

**NAUDAN SORKKA-ALUEEN IHOTULEHDUKSEN  
BAKTERIOLOGINEN TUTKIMUS 16S rRNA-  
RINNAKKAISSKVENSOINNILLA**

Tuomas P Hintikka  
Pro gradu -tutkielma  
Maatalous-metsätieteellinen  
tiedekunta  
Helsingin Yliopisto  
Mikrobiologian laitos, 2/2020



Tiedekunta – Fakultet – Faculty Maatalous-metsätieteellinen tiedekunta		Koulutusohjelma – Utbildningsprogram – Degree Programme - Mikrobiologia
Tekijä – Författare – Author Hintikka Tuomas P		
Työn nimi – Arbetets titel – Title NAUDAN SORKKA-ALUEEN IHOTULEHDUKSEN BAKTERIOLOGINEN TUTKIMUS 16S rRNA- RINNAKKAISSEKVENSOINNILLA		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track – Mikrobiologia - Microbiology		
Työn laji – Arbetets art – Level Pro gradu -tutkielma – Master's thesis	Aika – Datum – Month and year 2/2020	Sivumäärä – Sidoantal – Number of pages 75 s. - 75 p.
<p>Tiivistelmä – Referat – Abstract</p> <p>Nautojen (<i>Bos taurus</i>) sorkka-alueen ihotulehdus (DD) on ympäri maailman levinnyt sairaus, jota on myös havaittu muilla sorkkaeläimillä. Nykytiedon mukaan nautojen DD on bakteeri-, ei virus- tai sienitauti. Se aiheuttaa haavoja, vaurioita ja känsiä. DD leviää helposti ja aiheuttaa eläimelle myös kipua, heikentää sen yleiskuntoa ja hyvinvointia sekä aiheuttaa tuotantotappioita. Keski-Euroopassa ja Yhdysvalloissa DD:n on havaittu leviävän lähes epidemialuonteisesti, kun taas Pohjoismaissa sitä on tavattu yksittäisissä karjoissa. Taudin etiologiasta, esiintyvyydestä ja parhaasta mahdollisesta hoitomuodosta on eri teorioita.</p> <p>Nykytiedon ja tieteellisen kirjallisuuden mukaan merkittävimpinä DD:n aiheuttajabakteereina pidetään eri <i>Treponema</i>-bakteerilajeja. Uusimpien tutkimusten perusteella taudin vakavuuteen vaikuttaa eri <i>Treponema</i>-lajien yhteisvaikutus keskenään tai muiden bakteereiden kanssa.</p> <p>Tämän tutkielman tavoite oli pystyttää bakteerien 16S rRNA-geenin rinnakkaissekvensointiin perustuva metataksonomiatyöväline bakteereiden tunnistamiseksi. Tutkielmassa vertailtiin kahta eri DNA-eristysmenetelmää, kahta sekvensointityövälinettä ja kahta bioinformatiikan analyysimenetelmää. Näyttemateriaalina oli suomalaisten nautojen sorkka-alueen ihosta otetut biopsianäytteet (n=10). Viisi näytteistä oli terveestä M0-naudasta ja viisi M2-naudasta.</p> <p>Tuloksena havaittiin pieniä määriä <i>Spirochaetaceae</i>-bakteeriperheen sekvenssejä kahdesta M0 -diagnosoduista naudasta. Kaikista M2 -diagnoosin naudoista havaittiin runsaasti <i>Spirochaetaceae</i> - sekvenssejä. Lisäselvityksiä <i>Treponema</i>-bakteerien luotettavasta lajitason tunnistamisesta sekä niiden roolista taudissa tarvitaan.</p> <p>Bovine (<i>Bos taurus</i>) digital dermatitis (DD) of the foot is a widespread disease around the world, and it has also been diagnosed with several other hoofed ruminants. According to current knowledge, bovine digital dermatitis is a bacterial disease, not a viral or fungal disease. It causes ulcers, lesions, and raised calluses. DD spreads easily, and it can cause pain to the animal, weaken its overall health and well-being, as well as cause loss of production. DD has been almost epidemic in Central Europe, as well as in the United States of America, whereas in Scandinavia it has only been encountered in individual herds.</p> <p>According to current knowledge and the scientific literature, the most important bacteria responsible for DD are <i>Treponema</i> species. Recent studies suggest that the severity of the disease is affected by the interaction of different <i>Treponema</i> species with each other or with other bacteria.</p> <p>The aim of this thesis was to set up a metataxonomic pipeline for the parallel sequencing of the 16S rRNA bacterial gene, in order to identify the bacteria. Two different DNA extraction methods, two different sequencing pipelines, and two different bioinformatics pipelines were evaluated in this thesis. Sample material was biopsy samples of the skin of the Finnish bovine hoof (n=10). Five of the samples were from healthy M0 cattle and five from M2 cattle (acute active ulcerative lesions).</p> <p>As a result, small amounts of <i>Spirochaetaceae</i> bacterial family sequences were detected in two M0 diagnostic cattle. Abundant <i>Spirochaetaceae</i> sequences were found in all bovine diagnosed with M2. Further studies on the reliable species identification of <i>Treponema</i> bacteria and their role in the disease are needed.</p>		
Avainsanat – Nyckelord – Keywords metataksonomia, 16S rRNA, nauta, digitaalinen dermatiitti, sorkka-alueen ihotulehdus, känsä, <i>Treponema</i>		
Ohjaaja tai ohjaajat –Handledare – Supervisor or supervisors Erikostutkija Taru Lienemann, Tutkimusprofessori Sinikka Pelkonen ja Professori Per Saris.		
Säilytyspaikka – Förvaringställe – Where deposited Mikrobiologian laitoksen kirjasto		

## Lyhenteet

dATP	Deoksiadenosiinitrifosfaatti
dCTP	deoksisytidiinitrifosfaatti
DD	Digitaalinen dermatiitti, nautojen sorkka-alueen ihotulehdus
ddNTPs	dideoksi-nukleotideja:  ddATP, ddGTP, ddCTP ja ddTTP
dGTP	deoksiguanosiinitrifosfaatti
DNA	Deoksiribonukleiinihappo
dTTP	deoksitymidiinitrifosfaatti
FISH	fluoresoiva koetin
ID	Interdigitaalinen nektobasilloosi, naudoilla sorkkavälin ajotulehdus
Metagenomi	Kaikkien solujen DNA:n selvittäminen määrätystä ympäristönäytteestä
MLST	Multilocus sequence typing, useiden eri geenialueiden tarkastelu
MOCK microbial population	Kontrollinäyte, jossa on bakteerien DNA:ta tarkasti määrätyissä suhteissa.

NGS	Yleisnimitys kaikkien Sanger-sekvensointitekniikan jälkeisille sekvensointitekniikoille.
PBS	Puskuriliuos, jonka osmolariteetti, ionikoostumus ja ionien konsentraatiot ovat samat kuin ihmiskehossa.
PCR	Polymeraasiketjureaktio, jossa entsyymillä muodostetaan alkuperäisestä DNA-juosteesta kopio
Ribosomi	Solun rakenne, jossa tapahtuu proteiinisynteesi
RNA	Ribonukleiinihappo
TBE	Tris/boraatti/EDTA-puskuriliuos
WGS	kokogenomisekvensointi
16S rRNA-geeni	Bakteereilla, arkeoneilla ja muilla yksisoluisilla erästä ribosomin erästä osaa koodaava geeni

## Luettelo taulukoista

<b>Taulukko 1.</b>	Spirokeetojen genomien ominaisuuksia.....	17
<b>Taulukko 2.</b>	Eri treponemalajeja ja niiden eristyslähteitä.....	17
<b>Taulukko 3.</b>	Naudan sorkka-alueen ihotulehduksen diagnoositerminologiaa.....	21
<b>Taulukko 4.</b>	ATCC MSA-4000-tuotteen koostumus.....	29
<b>Taulukko 5.</b>	Kaupallisen MetaPolzyme-entsyymisekoitteen sisältämät entsyymit.....	30
<b>Taulukko 6.</b>	Illuminan 16S rRNA - sekä Kozich ym. 2013 -työvuot ja niiden vertailua.....	34
<b>Taulukko 7.</b>	PCR-reaktion reaktioseoksen komponentit eri työvuokaavioissa.....	35
<b>Taulukko 8.</b>	Illuminan työvuon PCR-ohjelmat.....	35
<b>Taulukko 9.</b>	Lopullinen Kozichin menetelmän PCR-ohjelma.....	36
<b>Taulukko 10.</b>	MiSeqin laatuparametreja tämän työn sekvensointiajoissa.....	40
<b>Taulukko 11.</b>	Sekvenssien kokonaismäärät eri sekvensoinneissa Chipster-laskennan perusteella.....	41

## Luettelo kuvista

<b>Kuva 1.</b> Illuminan MiSeq-sekvensaattori.....	13
<b>Kuva 2.</b> Vasikan sorkka halkileikattuna.....	22
<b>Kuva 3A ja 3B.</b> Erikoiskoulutettu eläinlääkäri tekemässä biopsianäytteenottoa puudutetusta naudan sorkasta..	26
<b>Kuva 4.</b> Tutkielman sovellettu työvuo Kozich ym. 2013 julkaisun pohjalta.....	28
<b>Kuva 5.</b> 16S rRNA sekvensointitulos, Chipster-laskenta, DNA-eristys Blood and Tissue kit.....	42
<b>Kuva 6.</b> 16S rRNA sekvensointitulos, Chipster-laskenta, näytteen kuulajahajotus, DNA-eristys Blood and Tissue kit.....	43
<b>Kuva 7.</b> 16S rRNA sekvensointitulos, Chipster-laskenta, DNA-eristysmenetelmä QuickPick SML gDNA.....	44
<b>Kuva 8.</b> 16S rRNA sekvensointitulos, Illumina 16S Metagenomics-laskenta, DNA-eristysmenetelmä QuickPick SML gDNA.....	45
<b>Kuva 9.</b> Kontrollien 16S rRNA sekvensointitulos, Chipster-laskenta, DNA- eristysmenetelmä QuickPick SML gDNA.....	46

# SISÄLLYSLUETTELO

Lyhenteet

Luettelo taulukoista

Luettelo kuvista

1. JOHDANTO.....	9
2. KIRJALLISUUSKATSAUS.....	11
2.1. Katsaus sekvensointimenetelmiin.....	11
2.2. Illuminan rinnakkaissekvensointitekniologia.....	12
2.3. 16S rRNA-geenin tutkiminen rinnakkaissekvensoinnilla.....	13
2.4. Sekvensointi ja naudan sorkka-alueen ihotulehdus.....	14
2.5. <i>Treponema</i> -suvun bakteereista ja niiden genetiikkaa.....	15
2.6. Nautojen sorkkasairauksista.....	18
2.6.1. Naudan sorkka-alueen ihotulehdus.....	18
2.6.2. <i>Treponema</i> -bakteerit naudoilla.....	19
2.6.3. Naudan elimistön puolustusreaktioita.....	20
2.6.4. DD:n diagnosointi.....	21
2.6.5. Nautojen DD:n nykyhoitosuositus Yhdysvalloissa.....	22
2.7. Bioinformatiikka.....	22
2.7.1. Eri bioinformaattisia algoritmeja.....	23
2.7.2. Chipster-ohjelmapaketti.....	23
2.7.3. Bakteereiden sekvenssitietokantoja.....	24
3. TUTKIELMAN TAVOITTEET.....	24
4. AINEISTO JA MENETELMÄT.....	24
4.1. Tutkielman näytemateriaali.....	24
4.2. Kontrollinäytteet.....	26
4.3. DNA-eristys sorkkabiopsianäytteistä.....	29
4.3.1. Näytteen esihajotus molemmilla kiteillä.....	30
4.3.2. Qiagen DNEasy Blood and Tissue -kitti.....	31
4.3.2.1. Alkuinkubointikokeilu Blood and Tissue -kitillä.....	31
4.3.2.2. Kuulahajotus- ja alkuinkubointikokeilu Blood and Tissue -kitin yhteydessä.....	31
4.4. DNA-eristys Bionobile Oy:n QuickPick SML gDNA -kitillä.....	32
4.5. DNA:n pitoisuuden ja puhtauden arviointi.....	32
4.6. Käytetyt sekvensointityövuot.....	32
4.6.1. Illuminan 16S rRNA-geenin tunnistamisen sekvensointityövuot.....	33
4.6.2. Kozich ym. 2013-julkaisuun perustuva 16S rRNA-geenin sekvensointityövuot.....	33
4.6.3. PCR-reaktiot eri sekvensointityövuokaavioissa.....	34

4.6.4. Geelielektroforeesi ja geelieristys.....	36
4.6.5. Geelieristys Gel Extraction -kitillä.....	37
4.6.6. Kirjaston valmistus Illuminan työvuolla.....	37
4.6.7. Kirjaston valmistus Kozichin työvuolla.....	37
4.6.8. MiSeqin latausliuos.....	38
4.7. Tutkielman bioinformatiikkamenetelmät.....	38
5. TULOKSET.....	39
5.1. Tulokset, 10 yleisintä bakteeritaksonia.....	39
5.1.1. Sekvensointi 19.11.2018.....	39
5.1.2. Sekvensointi 29.5.2019.....	39
5.2. Positiivisten kontrollien kontrolli.....	41
6. TULOSTEN TARKASTELU.....	47
6.1. Näytteenotto, näytteen käsittely ja nautojen hyvinvointi.....	47
6.2. DNA-eristys.....	48
6.2.1. Alkuinkuboinnin tarkastelua.....	48
6.2.2. Blood and Tissue kitin tarkastelua.....	49
6.2.3. Bionobile QuickPick SML gDNA kitin tarkastelua.....	49
6.2.4. Geelieristuksen tarkastelua.....	50
6.3. Sekvensointityövuokaavioiden tarkastelua.....	50
6.3.1. Sekvensoinnit V2 ja V3-tuotteilla.....	51
6.3.2. Alukkeet ja PCR-reaktiot.....	52
6.4. Kirjaston valmistus ja näytteen lataus.....	52
6.5. Qubit-, DeNovix ja Bioanalyzer-mittalaitteet.....	53
6.6. Bioinformatiikkaosuuden tarkastelua.....	53
6.7. Tulosten esitys ja rarefaktio.....	55
6.8. Bakteriologisten tulosten tarkastelu.....	55
6.8.1. Vertailu kuulajahajotuksen sekä normaalin DNA-eristuksen välillä...56	
6.8.2. Vertailu Chipster-laskennan ja 16S Metagenomics-laskennan välillä .....	57
6.8.3. Kontrollit.....	58
6.9. 16S rRNA-geenin sekvensoinnin käyttökelpoisuus, jatkotutkimukset...59	
7. JOHTOPÄÄTÖKSET.....	62
Huomiot ja kiitokset.....	65
LÄHDELUETTELO.....	67
Liite 1. Chipster-ohjelmopakettin 16S rRNA-tuloslaskennan eri työvaiheet..	74
Liite 2. Menetelmäkehitysskennon 19.11.2018 DNA- eristysparametreja.....	75
Liite 3. 29.5.2019 sekvensointiajon näytteet.....	76
Liite 4. Chipsterin 16S rRNA laskennan parametrit.....	77
Liite 5. <i>Bacteroidetes/Firmicutes</i> -suhteen tarkastelu.....	78



## 1. JOHDANTO

Nautojen (*Bos taurus*) sorkka-alueen ihotulehdus (digitaalinen dermatiitti, DD) on ympäri maailman levinnyt sairaus, jota on myös havaittu muilla sorkkaeläimillä. DD:n tuntomerkkejä ovat naudan sorkka-alueen ihon haavat, vauriot tai känsät. Nykytiedon mukaan nautojen DD:lla on bakteriologinen tausta ja tieteellisen kirjallisuuden mukaan merkittävimpinä DD:n aiheuttajina pidetään eri *Treponema*-bakteerilajeja (Klitgaard ym., 2008; Peek & Divers, 2018). Virusten ja sienten mahdollista osuutta on tutkittu shotgun-sekvensoinnin avulla (Krull ym., 2014). DD leviää karjassa helposti ja se aiheuttaa eläimelle kipua, heikentää sen yleiskuntoa ja hyvinvointia sekä aiheuttaa tuotantotappioita (Peek & Divers, 2018). Keski-Euroopassa ja Yhdysvalloissa DD:n on havaittu leviävän lähes epidemialuonteisesti, kun taas Pohjoismaissa sitä on tavattu yksittäisissä karjoissa (Kontturi ym., 2019; Peek & Divers, 2018). Yhdysvalloissa taudin taloudellisiksi tappioiksi on arvioitu noin 190 miljoonaa dollaria vuodessa, toisen tutkimuksen mukaan taas vähintään 130 USD nautaa kohden. Iso-Britanniassa taudin on arvioitu maksavan 99 puntaa eläintä kohden ja Alankomaissa 65 nautaa kohden arviolta 1517 USD vuodessa (Evans ym., 2016).

Bakteerigeenien sekvensointimenetelmä on melko uusi teknologia, ja uuden sukupolven sekvensointilaitteistolla on pystytty tutkimaan vuosikymmeniä tai -satoja vanhoista näytteistä esimerkiksi *Yersinia pestis* -bakteerin geenejä ja vertaamaan niitä uudempiin saman bakteerin näytteisiin (Bos ym., 2016). Bakteerigeenien sekvensoinnilla on osoitettu myös bakteerin maantieteellistä kulkeutumista ja muuntumista nautapopulaatiossa (Dupuy ym., 2012).

Tämä pro gradu -tutkielma on osa maa- ja metsätalousministeriön koordinoimaa "Lypsylehmien sorkka-alueen ihotulehdus -uusi uhka lypsykarjataloudelle" -hanketta, johon osallistuivat myös Valio Oy, Eläinten terveys ETT ry, Helsingin yliopiston eläinlääketieteellinen tiedekunta ja Ruokavirasto. Tutkielman tuloksilla voi pitkällä aikavälillä tarkasteltuna olla vaikutusta nautaeläinten terveellisempään ja parempaan elämänlaatuun.

## 2. KIRJALLISUUSKATSAUS

### 2.1. Katsaus sekvensointimenetelmiin

Biologisissa tieteissä sekvensoinnilla tarkoitetaan DNA-molekyylin rakenteen tarkan nukleotidijärjestyksen selvittämistä. Uuden sukupolven sekvensointitekniikat (engl. Next Generation Sequencing, NGS) ovat mahdollistaneet sekvensoinnin nopeuttamisen ja tehneet menetelmästä kustannustehokkaamman. Useiden DNA-kappaleiden sekvensointi samanaikaisesti samassa reaktiossa mahdollistaa suuren määrän rinnakkaissekvensointeja. NGS on ollut yleisnimitys kaikille Sanger-sekvensoinnin jälkeen kehitetyille uusille sekvensointitekniikoille. Ensimmäisen sukupolven sekvensointi, ns. Sanger-sekvensointi perustuu entsymaattiseen, PCR:n kaltaiseen reaktioon, jossa monistetaan polymeraasientsyymillä, neljän deoksinukleotidin (dATP, dTTP, dCTP ja dGTP) ja fluoresenssileimattujen dideoksinukleotidien (ddNTPs) avulla sattumanvaraisesti eripituisia DNA-fragmentteja. Dideoksinukleotidit ovat muutoin samanlaisia kuin nukleotidit, mutta niiden 3'-päässä oleva hydroksyyli-ryhmä on korvattu vetyatomilla, jolloin polymeraasientsyymien toiminta pysähtyy. Fragmentit, joiden koko on yleensä n. 500-1000 nukleotidia, on mahdollista erotella fluoresoivien leimojen avulla koon perusteella ja saada näin selville eri nukleotidien järjestys DNA-juosteessa (Liu ym., 2012; Madigan ym., 2019). Sanger-sekvensointi on hidasta, mutta sitä käytetään kuitenkin yhä ympäri maailman.

Eri sekvensointijärjestelmiä ovat olleet toisen sekvensointisukupolven 454-pyrosekvensointi ja synteisiin perustuva sekvensointi (Illumina, Yhdysvallat), Applied Biosystems'in SOLiD, kolmannen sukupolven HeliScope Single Molecule Sequencer ja Pacific Biosciences SMRT. Neljännen sukupolven sekvensointijärjestelmiä ovat tähän mennessä Ion torrent

(Thermo Fisher Scientific, Yhdysvallat) ja kannettava sekvensaattori MiniION (Oxford Nanopore, Yhdysvallat). Yksittäisen solun tai molekyylin sekvensointiin näistä soveltuvat Oxford Nanopore, SMRT- ja HeliScope-järjestelmät, muut näistä teknologioista käyttävät näyte-DNA:n monistamista (amplification). Rinnakkaisten näytteiden samanaikaiseen sekvensointiin soveltuvat 454-pyrosekvensointi, SOLiD- ja Illuminan teknologiat (Madigan ym., 2019).

Sekvensoitavan juosteen tai fragmentin pituus vaihtelee eri teknologioiden välillä runsaasti. 454-pyrosekvensoinnissa fragmentin pituus on 700 emäsparia, Illuminan HiSeqissa 2x150 tai 2x250 emäsparia, SMRT:llä yli 50000, SOLiD-järjestelmässä 50+35 tai 50+50 emäsparia (Liu ym., 2012; Madigan ym., 2019). Neljännen sukupolven sekvensointijärjestelmät kehittyvät yhä, ja tällä hetkellä niiden virhe on noin 5-15% (Pollock ym., 2018).

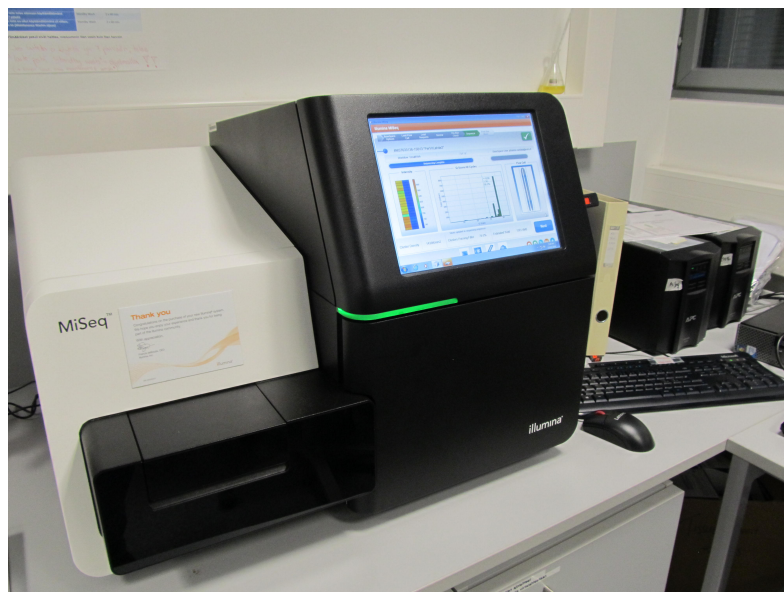
## **2.2. Illuminan rinnakkaissekvensointiteknologia**

Synteesiin perustuva sekvensointi on Illuminan patentoima sekvensointimenetelmä ja sen 16S rRNA -geenin tunnistamisen työväline koostuu useasta eri vaiheesta: geenispesifisillä alukkeilla tehdään määrämittäisiä kopiojuosteita alkuperäisestä DNA-juosteesta ja eri näytteiden kopiojuosteet yhdistetään yhdistelmänäytteeksi (geenikirjasto). Yhdistelmänäyte ladataan laitteeseen, jossa kopiojuosteet kiinnittyvät lukulevylle. Lukulevylle sekvenssiä luetaan sitä mukaa kun sekvensoitavalle DNA-juosteelle syntetisoidaan vastinjuostetta.

Jokainen eri näyte sisältää sille ainutkertaisen viivakoodin (engl. index), jonka avulla on mahdollista tunnistaa ja erotella eri näytteiden lukutulokset. Lukulevylle jokaiseen eri nukleotidiin (A,T,G,C) lisätään eri fluoresoiva väri, fluorofori, ja tämän jälkeen lukulevylle suunnataan laservalo. Fluoroforiin

osuva laserin valokvantti vapauttaa hetken päästä fluoroformimolekyylistä uuden valokvantin, ja tämän valokvantin erivärisyyden perusteella tarkastellaan näytteen nukleotidijärjestystä.

Tämän lisäksi laite tarkastelee lukutuloksia eri suodattimien sekä algoritmien avulla ja muuntaa ne näytekohtaisiksi tiedostoiksi (Kozich ym., 2013). Illuminan MiSeq-laitteella (Kuva 1) on mahdollista selvittää tietyn mikrobiomin, esimerkiksi sorkan ihon eliöyhteisön koostumusta (Klitgaard ym., 2014), sekvensoida näytteen RNA-molekyylejä tai usean eri bakteerin koko genomi samanaikaisesti (Madigan ym., 2019).



Kuva 1. Illuminan MiSeq-sekvensaattori.

### 2.3. 16S rRNA-geenin tutkiminen rinnakkaissekvensoinnilla

Tällä hetkellä yleinen bakteereiden metataksonominen tunnistamismenetelmä perustuu bakteereiden ribosomaalista RNA:ta koodaavan geenialueen tarkasteluun. Kyseiseltä geenialueelta on edelleen osoitettavissa monia eri alueita, esimerkiksi V1-V9, joista osa on ns. konservoituneita (eivät muutu ajan tai evoluution vaikutuksesta suurestikaan) ja osa taas on ns. hypervariaabeleita (muuttuvat runsaasti

ajan tai evoluution myötävaikutuksesta). Näiden geenialueiden sekvensointitulosten perusteella on tehty eri tietokantoja, joihin saatuja sekvensointituloksia voidaan verrata (Edgar, R.C., 2018; Kozich ym., 2013).

16S rRNA-geenin sekvensointimenetelmällä on etuja verrattuna perinteisiin bakteereiden viljelymenetelmiin, koska myös sellaisia bakteereita, arkeoneja ja muita yksisoluisia, joiden viljely ei onnistu perusmenetelmin, voidaan tunnistaa. Eläinsoluissa voi myös esiintyä solunsisäisiä, usein obligaatteja parasiitteja, joiden eristäminen ja kasvatus käytössä olevin viljelymenetelmin voi olla vaikeaa. Useat bakteerit myös vaativat muiden bakteereiden läsnäoloa elintoimintoihinsa. Bakteereiden 16S rRNA-geenialueen sekvensointimenetelmää on käytetty esimerkiksi kansainvälisessä Earth Microbiome projectissa bakteereiden yhtenä tunnistamismenetelmänä. Eri lajien evoluutiosta ja vuorovaikutuksesta keskenään voidaan tehdä hypoteeseja eri emäsjärjestyksen esiintymisen tai poissaolon perusteella. On myös huomattava, että eräillä bakteereilla voi myös olla monta kopiota samasta tarkasteltavasta geenialueesta genomissaan (Madigan ym., 2019).

#### **2.4. Sekvensointi ja naudan sorkka-alueen ihotulehdus**

Sekvensointia on käytetty myös nautojen sorkka-alueen ihotulehduksen mahdollisen mikrobiologisen riskitekijän selvitystyössä. Sorkan bakteriologisiin tutkimuksiin on käytetty mm. pyrosekvensointia (Klitgaard ym., 2013), bakteereiden 16S rRNA-geenin sekvensointia (Demirkan ym., 2018; Klitgaard ym., 2014) ja metagenomin selvittämiseen shotgun-sekvensoinnilla (Zinicola ym., 2015) tai myös samanaikaista 16S rRNA-rinnakkaissekvensoinnin ja shotgun-sekvensoinnin yhdistelmää (Krull ym., 2014).

## 2.5. *Treponema*-suvun bakteereista ja niiden genetiikkaa

*Treponema*-bakteerit ovat patogeeneja tai kommensaaleja (Madigan ym., 2019). Ne ovat laaja joukko, ja DSMZ:n hallinnoima bacterio.net-hakupalvelu listaa niitä 28 eri lajia (Bacterio.net-tietokanta, <http://www.bacterio.net/treponema.html> [Viittauspäivä 22.2.2020]). Bergey's Manual of Systematics of Archea and Bacteria listaa niitä 25 lajia (Norris ym., 2015) ja toinen tutkimus myös 25 eri lajia (Kuhnert ym., 2020). Ne ovat anaerobeja tai mikroaerofiilejä, korkkiruuvimaisia tai spiraalin kaltaisia liikkuvia bakteereita, joilla on flagella (Madigan ym., 2019; Norris ym., 2015). Ne ovat Gram-negatiivisia, 0,1-0,7 mikrometriä halkaisijaltaan ja pituudeltaan 1-20 mikrometriä (Norris ym., 2015). Morfologisena eroavaisuutena muista spirokeetoista *Treponema*-bakteerit ovat litteitä ja aaltomaisia (Madigan ym., 2019). *Treponema*-bakteerien viljely puhdasviljelmäksi on vaikeaa, joten jo perustutkimuksen ja -taksonomian muodostaminen on ollut haastavaa, mutta ne voidaan osoittaa suoramikroskopoinnilla (Klitgaard ym., 2014; Knauf ym., 2018).

*Treponema*-lajeista on tutkittu parhaiten ihmisillä esiintyviä eri lajeja ja alalajeja. Tropiikissa ja yleensä sen köyhimmillä alueilla ihmisillä esiintyvän vaapukkasyylätaudin on osoitettu aiheutuvan *Treponema pallidum* subsp. *pertenue* -bakteerin infektiosta. Ihmisille systeemisiä infektioita aiheuttaa *T. pallidum* subsp. *pallidum*, jonka genomi on vain noin 1,14 miljoonan emäsparin kokoinen, mikä on huomattavasti pienempi kuin esimerkiksi *T. denticolalla* (Taulukko 1).

Kokogenomisekvensoinnin perusteella on havaittu, että *Treponema*-bakteereiden genomi on ollut reduktiivisen evoluution kohteena, kun isäntänä on toiminut ihminen. Tämän teorian mukaan eri metaboliset reitit ovat karsiutuneet pois, koska isäntäeläimen kudospotilaalissa on ollut

saatavilla runsaasti eri ravinteita (Norris & Weinstock kirjassa Radolf, J.D., & Lukehart, S., 2006).

*Treponema phagedenis* -lajin eri kantojen 16S rRNA-geenin alue on hyvin konservoitunut, vastaavuus niiden välillä on yli 99%. Kahden ihmisistä ja kuuden naudoista eristetyt *Treponema*-kannan kokogenomisekvenssejä bioinformaattisesti verrattaessa havaittiin, että selvä fylogeneettinen ero oli osoitettavissa näiden kahdelta eri isännältä eristettyjen bakteereiden välillä. Tutkimuksessa tarkasteltujen, naudoilta eristettyjen bakteerikantojen genomit olivat suljettuja (Kuhnert ym., 2020). Kuhnertin ym. tutkimustulos myös vahvisti aiempaa MLST-analyysiin perustuvaa tutkimustulosta, jossa verrattiin 70 eri *Treponema phagedenis* -bakteerikannan geenialueita, kantojen eristyslähteet olivat hirvi, liha- ja lypsykarja, vuohi, lammas, sika ja ihminen. Tutkimuksessa havaittiin, että ihmisiltä eristettyjen kantojen genomeita keskenään verrattaessa niissä oli vaihtelua kuten eläimiltä eristetyissä kannoissa, mutta fylogeneettisen puun perusteella ne selvästi erosivat eläimiltä eristetyistä kannoista (Clegg ym., 2016). Taulukkoon 2 on koottu tietoa eri *Treponema*-bakteereista isäntineen.



**Taulukko 1.** Spirokeetojen genomien ominaisuuksia. Taulukko laadittu Seshadri ym. kirjassa Radolf, J.D., & Lukehart, S., 2006 mukaisesti.

Laji	Kromosomin muoto	Genomin koko	GC-pitoisuus %
<i>Treponema pallidum</i>	Yksittäinen, rengas	1,14 Mb	52,8
<i>Treponema denticola</i>	Yksittäinen, rengas	2,8 Mb	37,9
<i>Leptospira interrogans</i>	Kaksi rengasmaista	Yht. 4,7 Mb	36,0
<i>Borrelia burgdorferi</i>	Yksittäinen, lineaarinen	0,9 Mb	28,6

**Taulukko 2.** Eri treponemalajeja ja niiden eristyslähhteitä.

Taksoni	Isäntä	Bakteeri-näyte	Viite
<i>T.socranskii</i> subsp. <i>buccale</i> ja subsp. <i>paredis</i> <i>T.pectinovorum</i> <i>T.maltophilum</i> , <i>T.medium</i> , <i>T.denticola</i>	Ihminen ( <i>Homo sapiens</i> )	Plakki suun ientaskuissa	(Baron & Murray, 1999), <i>Bacterio.net-tietokanta</i> <a href="http://www.bacterio.net/treponema.html">www.bacterio.net/treponema.html</a> , [Viittauspäivä 22.2.2020].
<i>T.phagedenis</i> , <i>T.minutum</i>	Ihminen ( <i>Homo sapiens</i> )	Talirauhaseritegenitaalialueella	(Baron & Murray, 1999)
<i>T.pallidum</i> subsp. <i>pallidum</i>	Ihminen ( <i>Homo sapiens</i> )	Genitaalialueen haavauma	
<i>T. brennaborens</i>	Nauta	Leesioid DD-naudan jalassa	(Schrack, 1999)
<i>T.porcinum</i>	Sika	Uloste	(Nordhoff ym., 2005)

Eri nisäkäslajeista vuohilla on havaittu sorkkien digitaalista dermatiittia, ja tämänhetkisten tutkimustulosten perusteella varsinkin *Treponema*-suvun bakteereiden lajisto on ollut sairaissa eläimissä samankaltainen kuin naudoilla (Peek & Divers, 2018). *Treponema pallidumia* on osoitettu esimerkiksi Gambiassa ja Senegalissa vihermarakatteihin kuuluvalta *Chlorocebus sabaeus* -lajilta, Tansanian *Papio anubis* -oliivipaviaaneilta sekä mangabeihin kuuluvalta *Cercocebus atys* -lajilta Norsunluurannikolta (Knauf ym., 2018). *Treponema paraluiscuniculi* -bakteerin on todettu aiheuttavan kaniineille kanisyfilistä, ja tämän bakteerilajin ei ole todettu tarttuvan ihmiseen (Graves & Downes, 1981; Šmajš ym., 2012). Tosin ei ole selvää, onko *T.paraluiscuniculi* erillinen laji (Kuhnert ym., 2020).

## **2.6. Nautojen sorkkasairauksista**

### **2.6.1.Naudan sorkka-alueen ihotulehdus**

Naudan DD tunnetaan myös mansikkajalka-, digitaalinen syylä-, interdigitaalinen papillomatoosi- ja digitaalinen dermatiitti-nimillä, suomeksi sen nimi on sorkka-alueen ihotulehdus. Sairaus sisältää useimmiten naudan elimistön immunologisen reaktion, akuutin tulehduksen. Ainakin vuodesta 1994 lähtien se on levinnyt epidemioittain sekä Yhdysvalloissa että Keski-Euroopassa. Taudin aikana esiintyy kudokseen kertynyttä tulehdusnestettä ja ihon dermis-osan solujen koon kasvua. Hypertrofiaa ja muutoksia on nähtävissä myös histopatologisesti (Döpfer ym., 1997; Klitgaard ym. 2008, 2014). Taudin aiheuttama naudan epänormaali seisoma-asento voi käydä ilmi myös sorkkien asymmetrisenä tai epänormaalina kulumisena. Tautia esiintyy Yhdysvalloissa enemmän lypsykarjalla kuin lihakarjalla, useimmiten ensimmäistä laktaatiota läpikäyvissä eläimissä, mikäli karjassa esiintyy tautia (Peek & Divers, 2018).

Sairaita nautoja tutkineet eläinpatologit ovat esittäneet sairaista eläimistä otettujen, hopeavärjättyjen histopatologisten leikkeiden tarkastelun perusteella eri spirokeetabakteereita sorkka-alueen ihotulehduksen taudinaiheuttajiksi (Döpfer ym., 1997, Klitgaard ym., 2008)

### **2.6.2. *Treponema*-bakteerit naudoilla**

Yhdysvaltalaisen julkaisun mukaan primäärissä ja edelleen sekundäärisessä, syvemmässä infektiossa voivat eri *Treponema*-lajit olla osallisina. Karjoja tilatasolla tarkasteltaessa *Treponema*-lajisto on samankaltainen, mutta tilojen välisiä eroja voi esiintyä. Samoja organismeja voi löytyä myös naudan pötsin sisällöstä, ulosteista tai karjan iholta ja samanaikaista *Dichelobacter nodosus* -bakteerin infektiota voi myös esiintyä (Peek & Divers, 2018).

Tanskalaisten nautanäytteiden tarkastelun perusteella DD:n bakteriologisiksi aiheuttajiksi on ehdotettu eri spirokeetoja (Klitgaard ym., 2008) ja myös *Porphyromonas*-suvun bakteereita (Klitgaard ym., 2008). Myös muissa tutkimuksissa on sairailta eläimiltä havaittu *Porphyromonas*-suvun bakteereita (Döpfer ym., 1997; Krull ym., 2014; Sullivan ym., 2015). Fusobakteereita oli PCR-reaktioiden perusteella suurimmassa osassa sairaita nautoja ja saman tutkimuksen terveiltä naudoilta niitä löytyi hyvin vähän (Sullivan ym., 2015). Eräässä tutkimuksessa todettiin *T. Brennaborense* -bakteerin aiheuttavan digitaalista dermatiittia karjalla, engl. hairy foot warts (Strub ym., 2007).

Suomessa nautojen sorkka-alueen ihotulehdusta (DD) ja sorkkavälin ajotulehdusta (ID) tutkittaessa yleisiä bakteerilöydöksiä ovat olleet *Dichelobacter nodosus*, *Fusobacterium necrophorum*, *Porphyromonas levii*, *Prevotella melaninogenica*, *Trueperella pyogenes* sekä eri *Treponema*-suvun bakteerilajit (Pokki, 2015, Kontturi ym., 2019). *Treponema*-bakteereiden

esiintyvyys suomalaisten nautojen sorkissa, joilla on diagnosoitu sorkkavälin ajotulehdus, on osoitettu PCR-reaktioin. Ajotulehduksen osalta akuutin vaiheen naudoilla on havaittu terveitä nautoja suuremmat *F.necrophorum*- ja *D.nodosus* -bakteerien esiintyvyydet (Kontturi ym., 2019).

Nautojen digitaalista dermatiittiä epäiltäessä on myös tutkittu mahdollista virusten ja sienten osuutta taudinaiheuttana, mutta yhdysvaltalaisessa tutkimuksessa shotgun-sekvensointia käyttäen virus-DNA:ta oli vain 0,5% kaikista sekvensseistä, sieni-DNA:ta sitäkin vähemmän. Tutkimuksessa osoitettu virus-DNA koostui pääosin bakteriofagien DNA:sta (Krull ym., 2014). Eläinlääkäreiden Suomessakin havaitsemien, nautojen sorkissa havaittujen mahdollisten syylien tai känsien aiheuttaja on nykytiedon mukaan epäselvä.

### **2.6.3. Naudan elimistön puolustusreaktioita**

Naudan elimistön puolustusmekanismit myös bakteriologisia taudinaiheuttajia vastaan perustuvat pääpiirteissään fyysiseen esteeseen taudinaiheuttajien kulkeutumiselle, luontaiseen immunitettiin sekä hankittuun immunitettiin (Madigan ym., 2019; Tizard, 2009). Sorkka-alueella fyysisenä esteenä toimii ehjä iho ja luontaisena immunitettina toimivat naudan elimistön eri tulehdusreaktiot. Hankittu immunitetti vaihtelee myös eläinkohtaisesti. Hankittuun immunitettiin vaikuttavat suuresti nautaeläimen koko elinkaaren aikainen altistus eri tautia aiheuttaville sienille, bakteereille ja viruksille sekä alkueläimille (Döpfer ym., 1997; Tizard, 2009). Ruotsissa on kehitetty tankkimaitonäytteille käytettävä, *Treponema*-bakteerien vasta-aineisiin perustuva ELISA-koe (Frössling ym., 2018).

#### 2.6.4. DD:n diagnosointi

Naudan sorkka-alueen ihotulehduksen diagnosoinnissa on käytetty viisi- tai kuusiportaista asteikkoa (Taulukko 3), asteikolla M0-M4 ja M4.1 (Berry ym., 2012; Döpfer ym., 1997; Saarela, 2017). Pohjoismaissa jo makroskooppisen tarkastelun perusteella on voitu erottaa monta eri sorkka-alueen muutosta, mm. anturahaavauma, valkoviivan paise, sorkkavälin ihotulehdus ja sorkka-alueen ihotulehdus (Sorkka-atlas, 2013). Vasikan halkileikattu sorkka, joka on myös erittäin liikkuva kokonaisuus, on kuvassa 2.

**Taulukko 3.** Naudan sorkka-alueen ihotulehduksen diagnoositerminologiaa.

Diagnoosikoodi	Selite
M0	Normaali sorkka-alueen iho
M1	Leesioita, aikainen vaihe, granulomatoottisia, pyöreitä, vain epiteelin pinta on vaurioitunut.
M2	Ulseratiivisia alueita ruununrajan lähellä, leesiot alle 7 cm halkaisijaltaan, akuutti muutos, "strawberrylike"
M3	Leesiot ovat jo paranevia, rupikudoksella peittyviä haavaumia
M4	Kroonisiksi muuttuneet leesiot, ihon keratiinikerroksen liikakasvua tai se on paksuuntunut tai molemmat yht'aikaisesti.



Kuva 2. Vasikan halkileikattu sorkka.

#### 2.6.5. Nautojen DD:n nykyhoitosuositus Yhdysvalloissa

Hoitokeinona sorkkaan asennetaan erittäin ohut ja lyhyt puuvillaside, jonka alle annosteltavalla salisyylihappo-oksiteasykliiniyhdistelmällä on saatu hyvin nopeita kivunlievityksiä sekä hyviä hoitotuloksia. *Dichelobacter nodosus* -bakteerin aiheuttamaa infektiota epäiltäessä jodi, kuparisulfaatti tai sinkki ulkoisesti siveltyinä ovat nykyisin hoitomuotoina. Saman naudan sorkan uusintainfektio on yleistä. Rokotekokeiluiden onnistumistulokset ovat olleet hyvin harvinaisia. Hoitokeinojen tulisi olla ensisijaisesti humaaneja eikä sorkkien spontaania parantumista kannata jäädä odottamaan (Peek & Divers, 2018).

### 2.7. Bioinformatiikka

Bioinformaattinen työ alkaa sekvensaattorin siirtäessä sekvenssitiedostot Internet-yhteyden välityksellä pilvipalveluun, josta edelleen tiedon järjestely, säilytys ja julkaiseminen ovat mahdollisia (Liu ym., 2012). Bioinformaattisen osan osalta eri parametreille asetettavat kynnsarvot vaikuttavat huomattavasti tuloksen luotettavuuteen. Annotaatioiden

oikeellisuuteen ja luotettavuuteen vaikuttaa myös tietokannan valinta (Edgar, R.C., 2018).

### **2.7.1 Eri bioinformaattisia algoritmeja**

Bioinformatiikan menetelmiä valittaessa otetaan samalla myös kantaa eri algoritmeihin, joita eri ohjelmat käyttävät. Burrows-Wheeler-algoritmia käytetään sekvensaattorilta saadun sekvenssitiedon kompressointiin ja lisäkäsittelyyn. Needleman-Wunsch-algoritmia käytetään sekvenssien linjaukselle (alignment), ja sitä voidaan myös muunnella sekä single-end-että paired-end-sekvensseille. Tämä algoritmi nopeuttaa tietokonelaskentaa ja sisältää alunperin kolme eri algoritmia. Algoritmiin voi laittaa eri laskenta-arvoja eri aukkosakkosäännöille.

Smith-Waterman-algoritmi on jatkokehitelty Needleman-Wunsch-algoritmista ja siinä algoritmin aukkosakkosäännöt eroavat hieman Needleman-Wunsch-algoritmista, aukko laskentamatriisissa voi saada myös arvon nolla. Smith-Waterman-algoritmia käytetään sekvenssitiedon linjaukselle (Polanski & Kimmel, 2007).

### **2.7.2. Chipster-ohjelmapaketti**

Tässä tutkielmassa käytettiin Tieteellisen laskennan keskuksen CSC:n (Espoo, Suomi) kehittämän, avoimen lähdekoodin Chipster-ohjelmapaketin 16S rRNA-työvuota sekvensointitulosten analysointiin. Chipsterin käyttöliittymä on Internet-osoitteessa <https://chipster.rahtiapp.fi/home>, [Viittauspäivä 10.2.2020]), ja käyttöliittymän taustalla toimii monta eri ohjelmaa sekä 16S rRNA -sekvenssidatalle sovellettua algoritmiä. Chipsterin 16S rRNA-työvuoto (Liite 1) on pääpiirteissään Mothur-tietokoneohjelman algoritmeihin perustuva, mutta siinä on myös osia muista ohjelmista, mm. VSEARCH-ohjelmasta (Rognes ym., 2016). Chipster käyttää SILVA-tietokantaa 16S rRNA-tulosten linjaukseen (Kallio ym., 2011).

### 2.7.3. Bakteereiden sekvenssitietokantoja.

Tuloksen analysoinnissa bioinformatiikkaohjelmistojen eri vaiheet läpikäyneitä sekvensointituloksia verrataan eri tietokantoja vastaan. Bakteerien geenisekvensointitietoa sisältäviä tietokantoja ovat mm. saksalainen SILVA (Quast ym., 2013), Michigan State Universityn RDP (Cole ym., 2014), yhdysvaltalainen National Institute of Healthin BLAST (Madden, 2002) ja yhdysvaltalainen Greengenes (DeSantis ym., 2006). Greengenes-tietokannasta on saatavilla vain vuoden 2013 versio avoimena, joten tutkimuskäyttöön soveltuvat SILVA, RDP ja BLAST-tietokannat. Eurooppalainen SILVA on kuratoitu ja kimeratarkistettu tietokanta (Quast ym., 2013).

## 3. TUTKIELMAN TAVOITTEET

Tutkielman tarkoituksena oli

- i) pystyttää toimiva bakteereiden DNA-eristys- ja 16S rRNA rinnakkaissekvensointimenetelmä naudan sorkan ihoalueen biopsianäytteille.
- ii) selvittää miten terveiden ja sorkan ihotulehdusta sairastavien nautojen sorkkavälin bakteerilajisto eroavat toisistaan
- iii) tutkia löytyykö sorkkanäytteistä *Treponema*-suvun bakteereja

## 4. AINEISTO JA MENETELMÄT

### 4.1. Tutkielman näyttemateriaali

Biopsianäytteet olivat maa- ja metsätalousministeriön Makera-hankkeen ”Lypsylehmiä sorkka-alueen ihotulehdus -uusi uhka lypsykarjataloudelle” -



hankekarjoista. Näytteenotolle oli eläinten omistajien suostumus. Näytteet oli kerätty ympäri Suomea sorkkahoidon yhteydessä ja näytteet sokkoutettiin HBAK-numeroin.

Menetelmäkehitysssekvensointiin 19.11.2019 valittiin 10 biopsianäytettä, joista tehtiin myös kuulajahotuksella käsitelty rinnakkaisnäytteet, jotka merkittiin tunnuksella B. Lisäksi ajoon otettiin vesikontrolli ja kolme eri reagenssikontrollia (Liite 2).

Sekvensointiin 29.5.2019, joka on tutkielman päätulos, valittiin 10 biopsianäytettä kahdeksalta eri tilalta sekä kontrollinäytteet. Viisi nautaa oli ryhmästä M0 (terve) ja viisi ryhmästä M2 (akuutti muutos) (Liite 3). Tähän sekvensointiin otettiin lisäksi mukaan 20 satunnaista nautanäytettä sekvensaattorin toiminnan parantamiseksi, kaksi vesikontrollia, kaksi positiivista kontrollia (ATCC MSA-4000 ja DSM 1788) ja kaksi kokeellista harjoittelukontrollia eri bakteeripuhdasviljelmistä. Illuminan menetelmän kokeilun osalta ensimmäiseen sekvensointiin otettiin yhteensä 60 nautanäytettä, jotka olivat jakautuivat M0-M4.1.-luokkiin, lisänä oli seitsemän kontrollia. Illuminan menetelmän toiseen sekvensointiin otettiin 21 nautanäytettä luokista M0-M4.1. ja yksi kontrolli.

Näytteenoton tekivät eläinlääkärit, joilla oli eläinkokeiden suorittamiseen ja koe-eläinten käsittelyyn vaadittava lisäkoulutus. Hankkeen näytteenotolla oli koe-eläinlupa. Hanke-eläinlääkäri tutki naudan sorkkaterveyden ja luokitteli sorkat diagnoosiluokkiin M0-M4.1. Naudan jalan annettiin puutua 10 minuutin ajan, näytteenottokohta pestiin viileällä vedellä ja tarpeen vaatiessa naudan jalka pestiin pehmeällä harjalla. Biopsianäyte otettiin halkaisijaltaan 6mm terällä Kruise Biopsy Punch-biopsianäytteenottimella. Näytteenoton jälkeen jalka suojattiin ja lääkittiin. Biopsianäyte siirrettiin steriiliin näyteputkeen (Sarstedt AG&Co. KG, Nümbrecht, Saksa), joka oli

täytetty RNAlater-liuoksella (Thermo Fisher Scientific Baltics, UAB, Vilna, Liettua). Näyteputki säilytettiin 1 vrk ajan jääkaapissa +4°C lämpötilassa, jonka jälkeen näytteet siirrettiin -20°C lämpötilaan myös kuljetuksen ajaksi. RNAlateria oli vähintään kymmenkertaisesti biopsianäytteen tilavuuteen verraten.

Tutkimushankkeen ja tutkielman näytteenottoa esitellään kuvassa 3A ja 3B.

Tutkielman menetelmäkaavio on kuvassa 4.



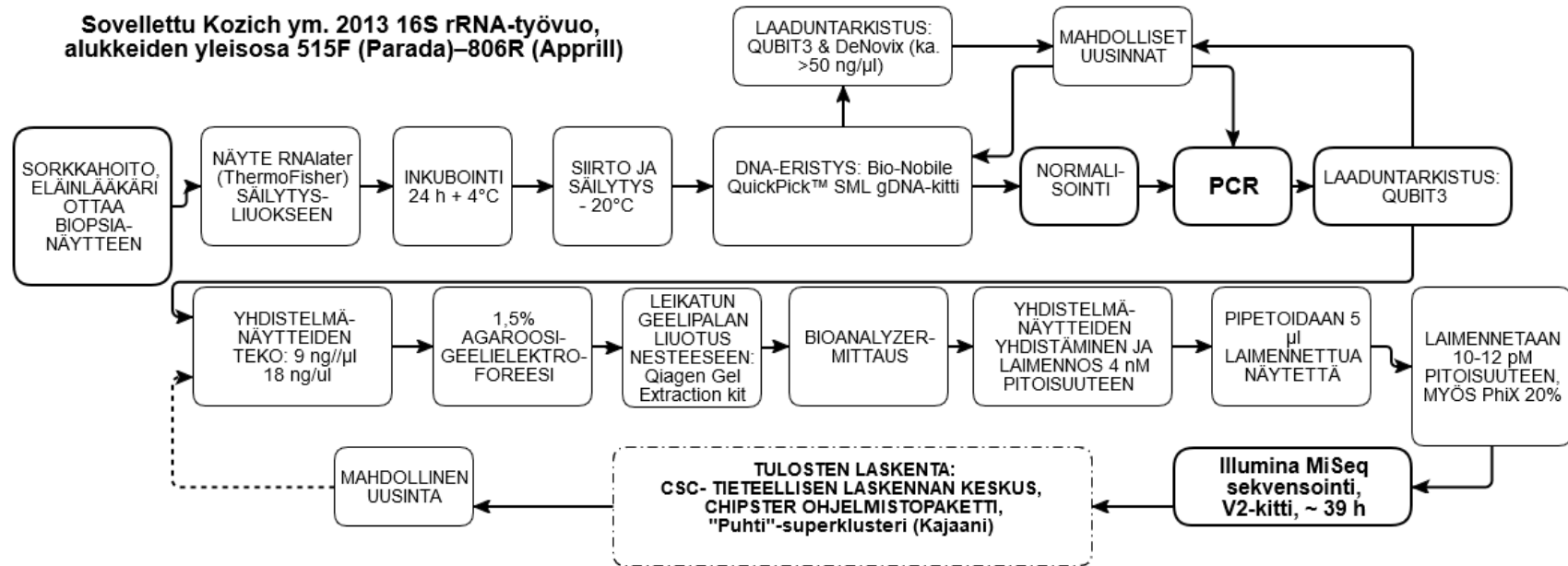
Kuvat 3A (vasemmalla) ja 3B (ylhäällä). Tohtorikoulutettava eläinlääkäri tekemässä biopsianäytteenottoa puudutetusta naudan sorkasta.

#### 4.2. Kontrollinäytteet

Näytteenotosta bioinformaattisiin tuloksiin asti työvuon läpi kuljetettavilla kontrollinäytteillä saadaan työvuon luotettavuutta arvioitua. Näitä ovat esimerkiksi reagenssikontrollit, tutkittavien bakteerien puhdasviljelmistä eristettyä DNA:ta sisältävät näytteet tai tarkasti määrätyissä suhteissa eri bakteerien DNA:ta sisältävät yhdistelmänäytteet (Hallmaier-Wacker ym.,

2018; Kozich ym., 2013; McGovern ym., 2018). Tutkielmassa negatiivisina kontrolleina käytettiin vettä ja reagensseja, joihin ei oltu laitettu DNA-templaattia työvuon eri vaiheissa.

Sekvensointiajossa käytettiin myös metataksonomiakontrolliksi soveltuvaa Metagenomic control material for Pathogen detection -bakteerisekoitusta (ATCC MSA-4000), joka sisälsi eri bakteereiden DNA:ta määrätyissä prosenttiosuuksissa (Taulukko 4), lisäksi kontrollina käytettiin saksalaisesta DSMZ-kantakokoelmasta *Treponema bryantii* (DSM 1788) -bakteerikannan puhdasta DNA:ta. Kontrolleja säilytettiin -80°C lämpötilassa.



Kuva 4. Tutkielman sovellettu työvuoto Kozich ym. 2013 julkaisun pohjalta.

**Taulukko 4.** ATCC MSA-4000-tuotteen koostumus.

0.1%	<i>Acinetobacter baumannii</i> strain 5377 (ATCC 17978)
0.7%	<i>Enterococcus faecalis</i> strain V583 (ATCC 700802)
1.4%	<i>Escherichia coli</i> strain CFT073(ATCC 700928)
14.4%	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> strain MGH78578 (ATCC 700721)
28.9%	<i>Neisseria meningitidis</i> strain FAM18 (ATCC 700532)
0.3%	<i>Pseudomonas aeruginosa</i> strain PAO1LAC (ATCC 47085)
0.7%	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> (MRSA) strain FPR3757 (ATCC BAA1556)
14.4%	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> (MSSA) strain TCH959 (ATCC BAA1718)
2.9%	<i>Streptococcus agalactiae</i> strain 2603 V/R (ATCC BAA611)
28.9%	<i>Streptococcus pneumoniae</i> strain Spain 23F1 (ATCC 700669)
7.2%	<i>Streptococcus pyogenes</i> strain SF370 (ATCC 700294)

#### 4.3. DNA-eristys sorkkabiopsianäytteistä

Tutkielmassa käytettiin kahta kaupallista kittiä DNA:n eristämiseen, Qiagen Blood and Tissue - ja Bio-Nobile Oy:n Quick Pick SML genomic DNA -kitit. Naudan sorkka-alueen kudosta sisältävä näytemateriaali on kirjallisuuden sekä pohjoismaisten tutkijoiden suullisten tietojen mukaan osoittautunut ongelmalliseksi. Molempien kittien yhteydessä käytettiin solujen hajotuksessa Metapolzyme-entsyymisekoitetta (Sigma-Aldrich Inc., Yhdysvallat), jossa on kuusi eri bakteereiden rakenteita hajottavaa entsyymiä (Taulukko 5). Entsyymiin lisättiin aseptisesti 750 µl PBS puskuria ja se varastoitui -20°C lämpötilassa.

Kummankin kitin yhteyteen lisättiin pidennetty, yön yli kestävä solujen hajotusvaihe. DNA-eristys sisälsi sekä entsyymaattisen että mekaanisen solujen hajotusvaiheen.

### 4.3.1. Näytteen esihajotus molemmilla kiteillä

DNA-eristystä varten sorkkanäytteet otettiin jääkaappiin sulamaan DNA-eristystä edeltävänä iltana. Biopsianäytteestä leikattiin steriilillä skalpellilla petrimaljan päällä 1/3-herneen kokoinen näytemassa, jota leikeltiin pienemmiksi paloiksi. Biopsianäytteestä pyrittiin saamaan ihoa ja vain hieman ihonalaista rustokudosta näytemassaan.

**Taulukko 5.** Kaupallisen MetaPolzyme-entsyymisekoitteen sisältämät entsyymit.

Entsyymi; alkuperäinen tuottoisäntä	Biokemiallinen vaikutuskohde	Valmistajan ilmoittama vaikutuskohde
Mutanolysiini (muramidaasi); <i>Streptomyces globisporus</i>	$\beta$ -N-asetyyli-muramyli-(1 $\rightarrow$ 4)-N- asetyyli-glukosamiinisidos bakteerin solunseinän peptidoglukaani-polysakkaridissa	Gram-positiiviset bakteerit
Akromopeptidaasi ( $\beta$ - lyyttinen proteaasi)	Gly-X-alueet peptidisekvensseissä, eritoten glysiini-glysiini-sekä D-Alaniini-X- alueet peptidoglykaaneissa	Gram-positiiviset aerobiset bakteerit
Lytikaasi; <i>Arthrobacter luteus</i>	Lineaaristen glukoosipolymeerien $\beta$ (1 $\rightarrow$ 3)- sidosten digestio	Hiivojen glukaani, aktiivisten hiivasolujen solunseinät
Kitinaasi; <i>Strepto- myces griseus</i>	Kitiinin entsyymaattinen hydrolyysi N-asetyyli-D- glukosamiiniksi	Kitiini
Lysostafiini; <i>Staphylo- coccus staphylolyticus</i>	Polyglysiini-crosslinks <i>Staphylococcus</i> - lajien solunseinässä	<i>Staphylococcus</i> - lajit, myös <i>S.aureus</i>
Lysotsyymi; Kananmunan valkuainen	$\beta$ (1-4)-sidosten hydrolyysi pepti- doglukaaniin	Gram- positiivisten bak- teerien peptidoglukaani- solunseinä

#### 4.3.2. Qiagen DNEasy Blood and Tissue -kitti

Ensimmäiset DNA-eristyskokeilut tehtiin pylväspuhdistukseen perustuvalla Qiagen Blood and Tissue-kitillä (Qiagen GmbH, Hilden, Saksa). Pilkottu näyte sisälsi runsaasti rusto- tai sarveiskudosta, joten kokeiluun valittiin kitin ohjeiden mukainen Rodent tails-menetelmä. Etanoli oli Etax Aa (Altia Oyj, Suomi).

##### 4.3.2.1. Alkuinkubointikokeilu Blood and Tissue -kitillä.

Alkuinkubointia varten näytemassa siirrettiin petrimaljalta autoklavoituun 2,0 ml Eppendorf Safe Lock -koeputkeen ja näytteeseen lisättiin 150 µl PBS-puskuria, 25 µl MetaPolyzyme-entsyymiä ja näytteitä inkuboituihin +35°C lämpötilassa lämpöhauteessa (Eppendorf ThermoMixer C, Eppendorf AG, Hampuri, Saksa) kuuden tunnin ajan. Tämän jälkeen lisättiin 180 µl ATL-liuosta, 20 µl proteinaasi K:ta ja inkuboituihin +56°C lämpöhauteessa yön yli, ravistus 300 rpm. Tämän jälkeen jatkettiin Rodent tails-menetelmällä.

##### 4.3.2.2. Kuulahajotus- ja alkuinkubointikokeilu Blood and Tissue -kitin yhteydessä

Rinnakkain samoille näytteille kokeiltiin myös kuulahajotusmenetelmää (beadbeating), jossa steriileillä, autoklavoiduilla kuulilla (PowerBead Tubes, glass 0,5 mm, Qiagen GmbH, Saksa) näytettä ravistaen hajotetaan näytekudosta eri aikoja. Näytemassa siirrettiin petrimaljalta kuulahajotusputkeen, johon pipetoitiin 400 µl PBS-puskuria. Kuulahajotusnäytteitä vorteksoitiin 60 tai 180 sekunnin ajan (Liite 2), jonka jälkeen nestefaasi ja mahdolliset sakkaumat siirrettiin puhtaaseen 1,5 ml Eppendorf-koeputkeen. Koeputkeen lisättiin kitin RNAasia 4 µl ja 180 µl ATL-puskuria ja laitettiin inkuboitumaan +35°C lämpöhauteeseen 6 h ajaksi. Näytteisiin lisättiin 20 µl proteinaasi K:ta, vorteksoitiin 10 sekunnin ajan ja näytteet laitettiin inkuboitumaan +56°C lämpötilaan yön yli (12 h). Tämän jälkeen eristystä jatkettiin Rodent tails -menetelmän mukaisesti.

#### **4.4. DNA-eristys Bionobile Oy:n QuickPick SML gDNA -kitillä**

Biopsianäytemassa siirrettiin petrimaljalta 2,0 ml Eppendorf-koeputkeen, johon oli pipetoitu 100 µl kitin lyysispuskuriliuosta. Näytteeseen pipetoitiin 10 µl Metapolyzyme-entsyymisekoitetta, vorteksoitiin kevyesti ja spinnattiin tarvittaessa. Eppendorf-putki laitettiin lämpöhauteeseen inkuboitumaan +35°C lämpötilaan 6 h ajaksi, ravistus 300 rpm. Tämän jälkeen lisättiin 20 µl proteinaasi K:ta, inkuboitiin +56°C lämpötilassa yön yli, kuitenkin enintään 16 tuntia, ravistus 300 rpm. Näyte vorteksoitiin, spinnattiin ja Binding Buffer-liuosta pipetoitiin näytteeseen 250 µl. Tämän jälkeen magneettipartikkeliliuoksesta pipetoitiin 8 µl magneettipartikkeleita ja Eppendorf-putki siirrettiin takaisin +56°C lämpöhauteeseen 8 min ajaksi, ravistelu 500 rpm. Eppendorf-putki siirrettiin huoneenlämpöön jäähtymään vähintään 30 minuutin ja enintään yhden tunnin ajaksi. Tämän jälkeen edettiin kitin ohjeen mukaisesti, pesupuskureiden tilavuus oli 500 µl ja ensimmäistä pesupuskuri Binding Bufferia lisättiin yksi työvaihe lisää. Lopullinen eluaatiotilavuus oli 100 µl ja magneettipartikkelit olivat eluaatiopuskurissa vähintään viiden minuutin ajan.

#### **4.5. DNA:n pitoisuuden ja puhtauden arviointi**

Eristetyn DNA:n pitoisuus mitattiin Qubit3-mittalaitteella (Invitrogen, Yhdysvallat) käyttäen Broad Range -kittiä. DNA:n puhtaus määritettiin DeNovix-mittalaitteella (DeNovix Inc., Wilmington, Yhdysvallat) A260/280 -aallonpituussuhteella. Kaikkien eristetyt DNA:t ylittivät tämän tutkielman laatuarvon  $A_{260}/A_{280} > 1,7$ .

#### **4.6. Käytetyt sekvensointityövuot**

Tutkielmassa käytettiin bakteerien tunnistamiseen kahta eri työvuota, joissa bakteerien 16S rRNA-geenialueen tunnistavat alukeparit sekä välivaiheiden määrät ovat erilaisia. Työvuokaavioiden eroja on käsitelty taulukossa 6.



#### **4.6.1. Illuminan 16S rRNA-geenin tunnistamisen sekvensointityövu**

Illuminan menetelmä perustuu kahteen eri PCR-reaktioon, jossa ensimmäisessä bakteerien 16S rRNA-geenin V3/V4-aluetta monistetaan PCR-reaktiossa yleisalukkeilla Bakt\_341F ja Bakt\_805R, joihin on lisätty Illuminan adapterit (Herlemann ym., 2011; Klindworth ym., 2013). Toisessa PCR-reaktiossa tuotteeseen lisätään yksilölliset indeksit. Työvuoto sisältää kaksi eri PCR-tuotteen puhdistusta esim. agarosigeeliltä tai esimerkiksi magneettikuulilla (Klindworth ym., 2013). Yleisalukkeet tilattiin MetaBion AG:lta (Saksa) ja indekseinä käytettiin Illuminan XT-indeksejä (Illumina, Yhdysvallat).

#### **4.6.2. Kozich ym. 2013-julkaisuun perustuva 16S rRNA-geenin sekvensointityövu**

Kozich ym. 2013-julkaisun työvuossa on vain yksi PCR-reaktio, koska indeksisuus muodostuu jo ensimmäisen PCR-reaktion aikana. Tutkielmassa käytettiin muokattuja B701-712 ja B501-B508-sarjan alukkeita, jotka oli päivitetty seuraavasti: Indeksiosa kuten Kozich ym. 2013, yleisosat olivat Earth Microbiome Projectin yleisalukkeet 515F (Parada ym., 2016) ja 806R (Apprill ym., 2015), read1-alukkeen muutos oli C<sub>16</sub>>Y, read2- alukkeen muutos oli H<sub>20</sub>>N ja index1-alukkeen muutos oli D<sub>13</sub>>N. Kaikki Kozichin menetelmän alukkeet tilattiin Metabion AG:ltä Saksasta, liuotettiin 100 µM pitoisuuteen TE-puskuriin ja käyttöä varten alukkeita laimennettiin PCR-vedellä (ThermoFisher Scientific, Vilna, Liettua) 5 nM konsentraatioon. Alukkeita säilytettiin -20°C lämpötilassa ja sulatettiin ennen käyttöä.

**Taulukko 6.** Illuminan 16S rRNA - sekä Kozich ym. 2013 -työvuot ja niiden vertailua.

<b>Työvaihe</b>	<b>Illumina 16S työvuoto</b>	<b>Kozich ym. 2013 työvuoto</b>
<b>DNA eristysmenetelmä</b>	Vapaasti valittavissa	Vapaasti valittavissa
<b>PCR</b>	2 kpl	1 kpl
<b>Sekvensoitavan tuotteen periaate</b>	Paired end	Paired end
<b>Monistuvan 16S rRNA-geenialueen koko</b>	~450 bp	~250 bp
<b>PCR-reaktiossa muodostuvan tuotteen lopullinen koko, emäsparia</b>	1. PCR ~ 460 2. PCR ~ 630	~ 420
<b>PCR, kierroksia yhteensä minimissään</b>	38	30
<b>Indeksointijärjestelmä</b>	Valmiit, kaupalliset XT-indeksit	Oma indeksointijärjestelmä sisällytettynä alukkeisiin
<b>Indeksien kiinnitys</b>	Indeksit kiinnitetään toisen PCR-reaktion aikana	Indeksiosuus muodostuu PCR-reaktion aikana

#### **4.6.3. PCR-reaktiot eri sekvensointityövuokaavioissa**

Tutkielman PCR-laitteina toimivat Biorad C1000 Touch Thermal cycler CFX-96 Real-Time (Biorad Inc., Hercules, Yhdysvallat) ja PTC-200 DNAEngine (MJ Research Inc., St. Bruno, Kanada). Polymeraasina ja reaktiosekoitteena molemmissa sekvensointityövuokaaviossa toimi Platinum SuperFI -valmiste (Invitrogen, Thermo Fisher Scientific, Vilna, Liettua). Kummassakin 16S-menetelmässä PCR-reaktiot tehtiin 200 µl:n PCR-putkiin. Kozichin työvuota optimoitiin GradientPCR sekä DNA-gradient-menetelmillä. Lopullisen PCR-

reaktioseoksen komponentit ja pitoisuudet ovat taulukossa 7 ja PCR-ajojen ohjelmat ovat taulukoissa 8 ja 9.

**Taulukko 7.** PCR-reaktion reaktioseoksen eri komponentit eri työvuokaaviossa.

Reagenssi	Illustrina 16S työvu, 1x reaktioseos	Pitoisuus reaktioseok- sessa	Kozich ym. 2013 työvu, 1x reaktioseos	Pitoisuus reaktioseok- sessa
Genominen DNA	2.5 µl	5 ng/µl	1,0 µl	5ng/µl
PCR-vesi	0 µl	-	9,5 µl	-
Aluke F	5 µl	1 µM	1,0 µl	5 µM
Aluke R	5 µl	1 µM	1 µl	5 µM
2xPlatinum SuperFI Mastermix	12,5 µl	1 x	12,5 µl	1 x

**Taulukko 8.** Illuminan työvuon PCR-ohjelmat.

	Illuminan 1. PCR		Illuminan 2. PCR	
Lämpötila	Aika		Aika	
95°C	0'30		0'30	
95°C	0'10	30 x	0'10	8 x
55°C	0'10		0'10	
72°C	0'45		0'45	
72°C	5'00		5'00	
4°C	∞		∞	

**Taulukko 9.** Lopullinen Kozichin menetelmän PCR-ohjelma.

<b>98°C</b>	0'30	
<b>98°C</b>	0'10	30 x
<b>57°C</b>	0'15	30 x
<b>72°C</b>	1'00	30 x
<b>72°C</b>	10'00	
<b>4°C</b>	∞	

#### 4.6.4. Geelielektroforeesi ja geelieristys

Molemmissa työvuokaavioissa PCR-reaktioiden toimivuutta tarkasteltiin agarosigeelielektroforeesin avulla (Sigma-Aldrich, Kalifornia, Yhdysvallat). Kaikki käytetyt geelit olivat 1,5% agarosigeelejä (Sigma-Aldrich, Yhdysvallat), ajopuskuri oli 1 x TBE-puskuri.

Agarosigeelielektroforeesin osalta käytettiin seuraavia reagensseja: värjäysaine oli SybrSafe Gel Stain (Invitrogen, USA), kokostandardina käytettiin GeneRuler 100 bp Plus DNA Ladder, 0,1 µg/µl (ThermoFisher, Vilna, Liettua), näytteen latauksessa käytettiin 6 x Loading Dye -latausväriainetta (ThermoFisher, Vilna, Liettua). Poikkeuksena tähän olivat Kozichin menetelmäkehityssekvensoinnin agarosigeelielektroforeesin reagenssit, jossa värjäysaineena oli GelRED Nucleid Acid Gel Stain -värjäysaine (Biotium, Yhdysvallat), kokostandardina toimi NEB Quickload 100 bp ja näytteen latauksessa oli New England Biolabs Loading Dye -latausväriaine (molemmat New England Biolabs, Yhdysvallat).

Geeliajon jännitteet olivat 100-120 V ja ajon kesto 60-75 minuuttia, laitteena oli Thermo Midicell (ThermoFisher Scientific, USA), ajopuskurina oli tuore 1 x TBE-puskuri. DNA:ta agarosigeeliltä eristettäessä kokostandardin mukaan määritettynä oikean kokoiset PCR-tuotteet leikattiin geeliltä steriilillä skalpellilla UV-valopöydän päällä ja siirrettiin 2,0 ml Eppendorf-putkeen.

#### **4.6.5. Geelieristys Gel Extraction -kitillä**

DNA eristettiin geelipaloista QIAquick Gel extraction -kitillä (Qiagen, Hilden, Saksa) seuraavin muutoksin kitin ohjeisiin: suoritettiin ylimääräinen Buffer QC-pesu, inkubaatioaika oli neljä minuuttia huoneenlämmössä ja eluaatiotilavuus oli 50 µl. Propanoli oli 2-propanolia (Sigma-Aldrich, Yhdysvallat).

#### **4.6.6. Kirjaston valmistus Illuminan työvuolla**

Illuminan työvuossa näytteet laimennettiin Tris-HCl + Tween20-liuoksella (pH 8,5 valm. pvm. 7.12.2018) Illuminan ohjeen mukaisesti 4 nM konsentraatioon. Tämän jälkeen tehtiin yhdistelmänäytteen normalisointi pipetoimalla 5 µl näytettä yhdistelmänäytteeseen.

#### **4.6.7. Kirjaston valmistus Kozichin työvuolla**

Kozichin työvuon ja menetelmäkehityksen osalta PCR-tuotteita näytteet normalisoitiin vasta ennen sekvensointia jakamalla ne kahteen eri yhdistelmänäytteeseen, joihin kustakin näytteestä pipetoitiin Qubit-mittausten perusteella sama määrä DNA:ta. Ensimmäisen yhdistelmänäytteen konsentraatio oli 9 ng/µl, johon pipetoitiin näytteet konsentraatioltaan 3-10 ng/µl. Toisen yhdistelmänäytteen konsentraatio oli 18 ng/µl, johon pipetoitiin näytteet, joiden DNA-konsentraatio oli suurempi kuin 10 ng/µl. Yhdistelmänäytteet pipetoitiin 1,5% agarosigeelille kahtena toistona. Geelille pipetoitavan yhdistelmänäytteen määrä oli 30 µl ja lisäksi 6 µl latauspuskuria, ajoaika oli 60 min ja jännite 100 V, ajolaitteena toimi Thermo Midicell. Tämän jälkeen ~420 bp kokoiset tuotteet leikattiin irti geeliltä, tehtiin DNA:n eristys geelistä Qiagen Gel Extraction kitillä ja tuotteiden pitoisuus mitattiin Bioanalyzer-laitteella (Agilent Technologies, Yhdysvallat) ja valittiin edustavimmat yhdistelmänäytteet sekvensointiin. Valitut yhdistelmänäytteet laimennettiin Qiaquick Gel extraction kitin eluaatiopuskurilla 4 nM

pitoisuuteen ja eri yhdistelmänäytteitä pipetoitiin sama määrä Eppendorf-putkeen.

#### **4.6.8. MiSeqin latausliuos**

Lopullinen MiSeqin latausliuos molemmissa työvuokaavioissa V2-kittiä käytettäessä sisälsi 10 pM näytekirjastoa ja 20% PhiX-kontrollifaagia (Illumina Inc., Yhdysvallat).

#### **4.7. Tutkielman bioinformatiikkamenetelmät**

Tulokset laskettiin CSC:n Chipster ohjelmapaketin 16S rRNA-työvuolla (Kallio ym., 2011) ja Chipster-laskennassa käytetyt parametrit on esitetty liitteessä 4. Lopullisen sekvensointiajon tulos laskettiin uudelleen "Max. number of ambiguous bases = 0" -arvolla.

Tuloksia laskettiin myös kaupallisella Illuminan Basespace-Internet-pilvipalvelun 16S Metagenomics-työkalulla, jonka erillisiä työvaiheita ei ole julkaistu ja jossa bakteerien tunnistus perustuu maksulliseen, suljettuun Greengenes-tietokantaan.

## 5. TULOKSET

Bakteeritaksonien osalta tulokset ovat kuvissa 5-8 ja kontrollien tulokset ovat kuvassa 9. Lisäksi esitetään MiSeqin laatuparametreja taulukossa 10 ja sekvenssien kokonaismäärät eri sekvensoinneissa (Taulukko 11).

### 5.1. Tulokset, 10 yleisintä bakteeritaksonia

Sekvensointituloksista poistettiin kaikki alle 500 sekvenssiä näytettä kohden sisältävät bakteeritulokset (annotaatiot) perustuen aiempiin vesikontrollien sekvensointituloksiin. Työvuoro oli Kozich ym., 2013 ja luvussa 4 esitetyn mukaisesti. Tulos esitetään prosentuaalisesti eri bakteeriperheiden osuutena näytteen kokonaissekvenssilukumäärän verraten.

#### 5.1.1. Sekvensointi 19.11.2018

Sekvensoinnissa näytteinä olivat normaalin DNA-eristyksen (Kuva 5) sekä kuulahajotuksen läpikäyneet (Kuva 6) biopsianäytteet. DNA-eristysmenetelmä oli Qiagen Blood and Tissue -kitti. Kuvista nähdään *Spichaetaceae*-perheen bakteereiden esiintymistä M4- tai M4.1.-naudoilla, joita saatiin osoitettua enemmän kuulahajotuksella hajotetuista näytteistä.

#### 5.1.2. Sekvensointi 29.5.2019

Sekvensoinnin biopsianäytteiden tulokset laskettiin Chipster-laskennalla (Kuva 7) ja Illuminan 16S Metagenomics -laskennalla (Kuva 8). DNA-eristysmenetelmä oli QuickPick SML gDNA -kitti. Lisäksi esitetään kontrollien tuloksia 29.5.2019 sekvensoinnissa (Kuva 9). Kuvista 7 ja 8 nähdään *Spirochaetaceae*-perheen bakteereiden runsas esiintyvyys M2-naudoilla.

Lisäksi liitteessä 5 on esitetty *Bacteroidetes/Firmicutes* -suhteen tarkastelu, kuvasta nähdään M0- ja M2-diagnoosin nautojen tulosten eroavan toisistaan.

**Taulukko 10.** MiSeqin laatuparametreja tämän työn sekvensointiajoissa.

	<b>selite</b>	<b>Kozich menetelmä- kehitys</b>	<b>Illumina 1-ajo</b>	<b>Illumina 2-ajo</b>	<b>Kozich 29.5.2019 ajo</b>
	<b>Clustering density [K/mm<sup>2</sup>]</b>	919 ± 22	350 ± 9	248 ± 2	537 ± 12
<b>R1</b>	<b>Ensimmäinen read, %≥Q30</b>	86.10	87.33	97.24	93.72
<b>R2</b>	<b>Ensimmäinen indeksi, %≥Q30</b>	69.46	30.97	19.41	31.80
<b>R3</b>	<b>Toinen indeksi,%≥Q30</b>	79.90	96.43	97.41	94.56
<b>R4</b>	<b>Toinen read, %≥Q30</b>	81.07	77.97	95.80	92.09

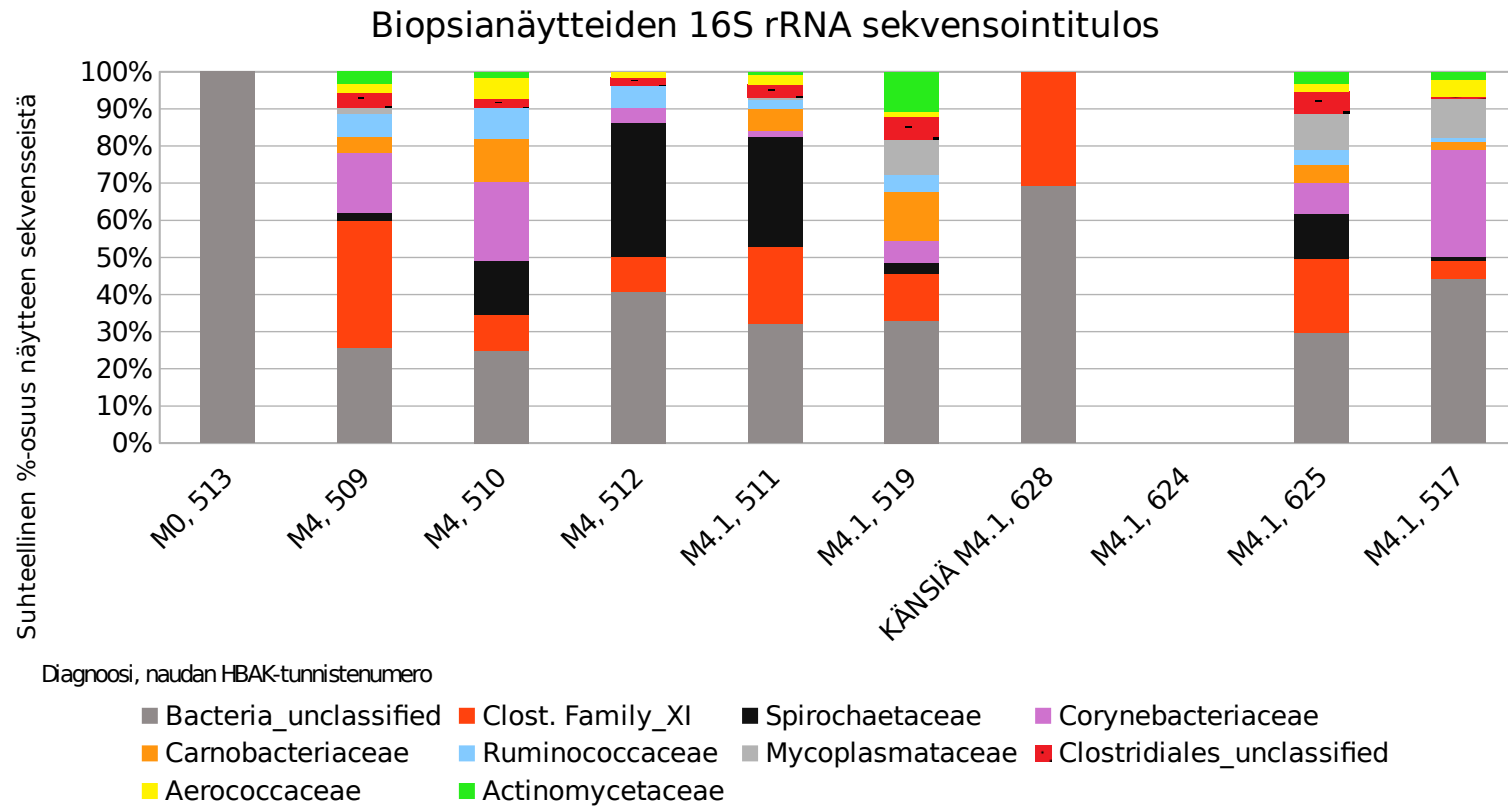


**Taulukko 11.** Sekvenssien kokonaismäärät eri sekvensoinneissa Chipster-laskennan perusteella.

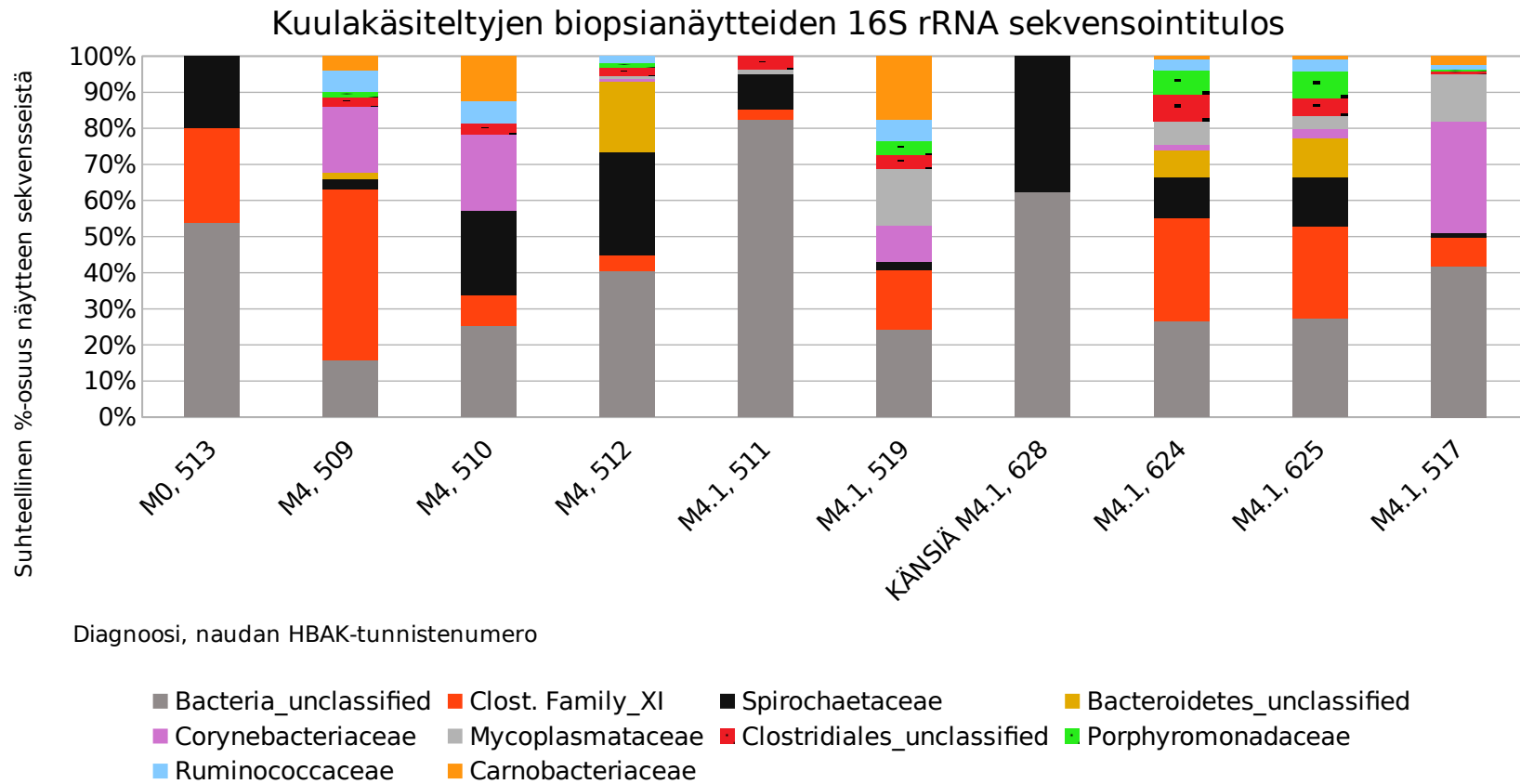
Sekvensointiajon nimi	Kozich menetelmä -kehitys	Illumina 1-ajo	Illumina 2-ajo	Kozich 2-ajo
Ajopäivämäärä	19.11.2018	19.3.2019	5.4.2019	29.5.2019
Työvuoto	Kozich ym. 2013	Illumina 16S	Illumina 16S	Kozich ym. 2013
Kitti	V2 (2x250 bp)	V3 (2x300 bp)	V2 (2x250 bp)	V2 (2x250 bp)
Näytteiden lukumäärä (sisältää kontrollit)	24	67	22	36
Esitiedostojen (engl. contig) lukumäärä, kpl sekvenssiä	2790716	580720	476474	2979383
Sekvenssejä kimeran poistamisen jälkeen	2690175	Ei täyttänyt laatuaroja	807	2699164
Uniikkeja sekvenssejä kimeran poistamisen jälkeen	562706	Ei täyttänyt laatuaroja	Ei tiedossa	98321

## 5.2. Positiivisten kontrollien kontrolli

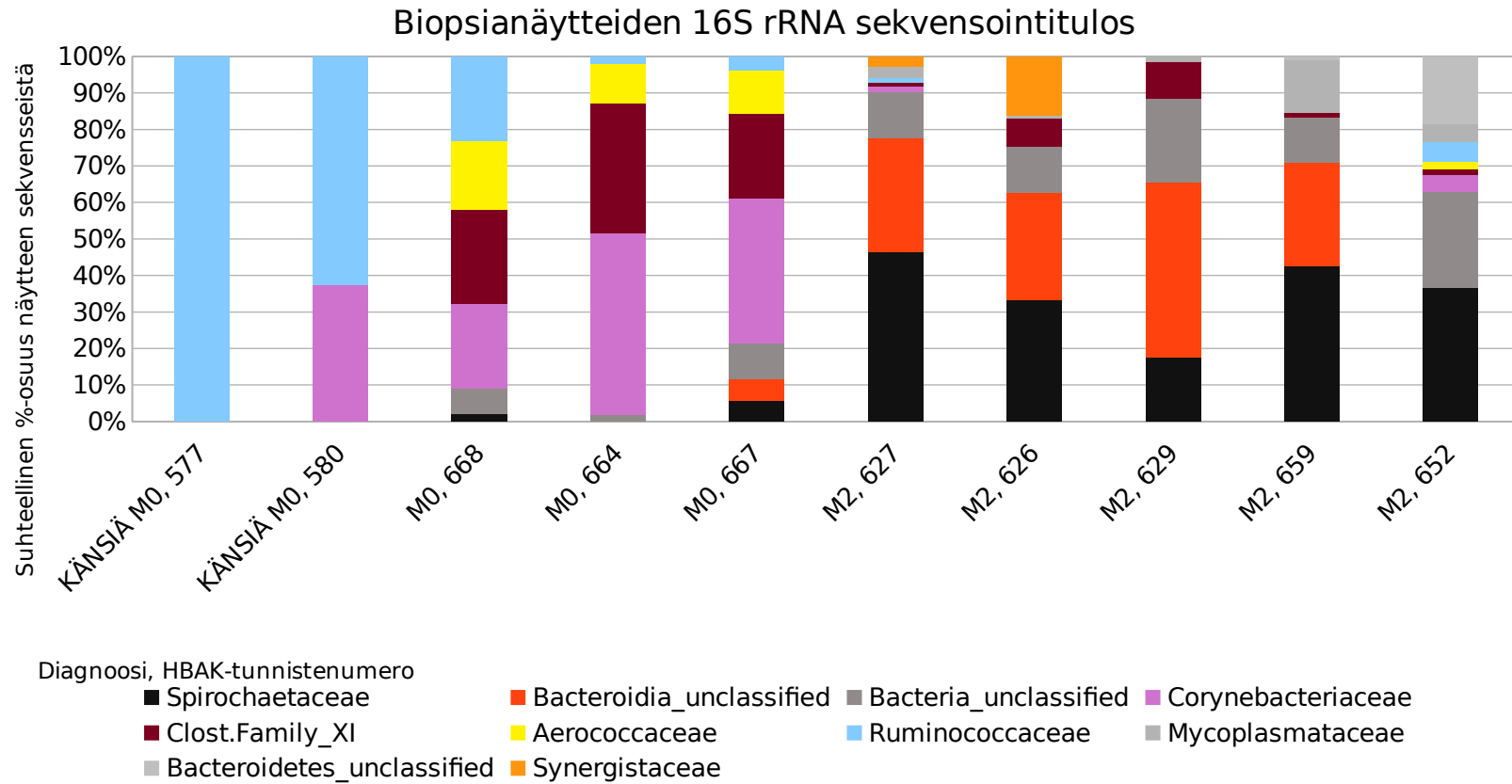
DSMZ:n ja ATCC:n positiivisia kontrolleja säilytettiin -80°C lämpötilassa. Laadunvarmistukseksi suoritettiin Qubit HS -mittaus kontrolleille, jonka perusteella ATCC:n MSA-4000 sisälsi 1,44 ng/μl DNA:ta, DSMZ:n kanta 19,8 ng/μl DNA:ta.



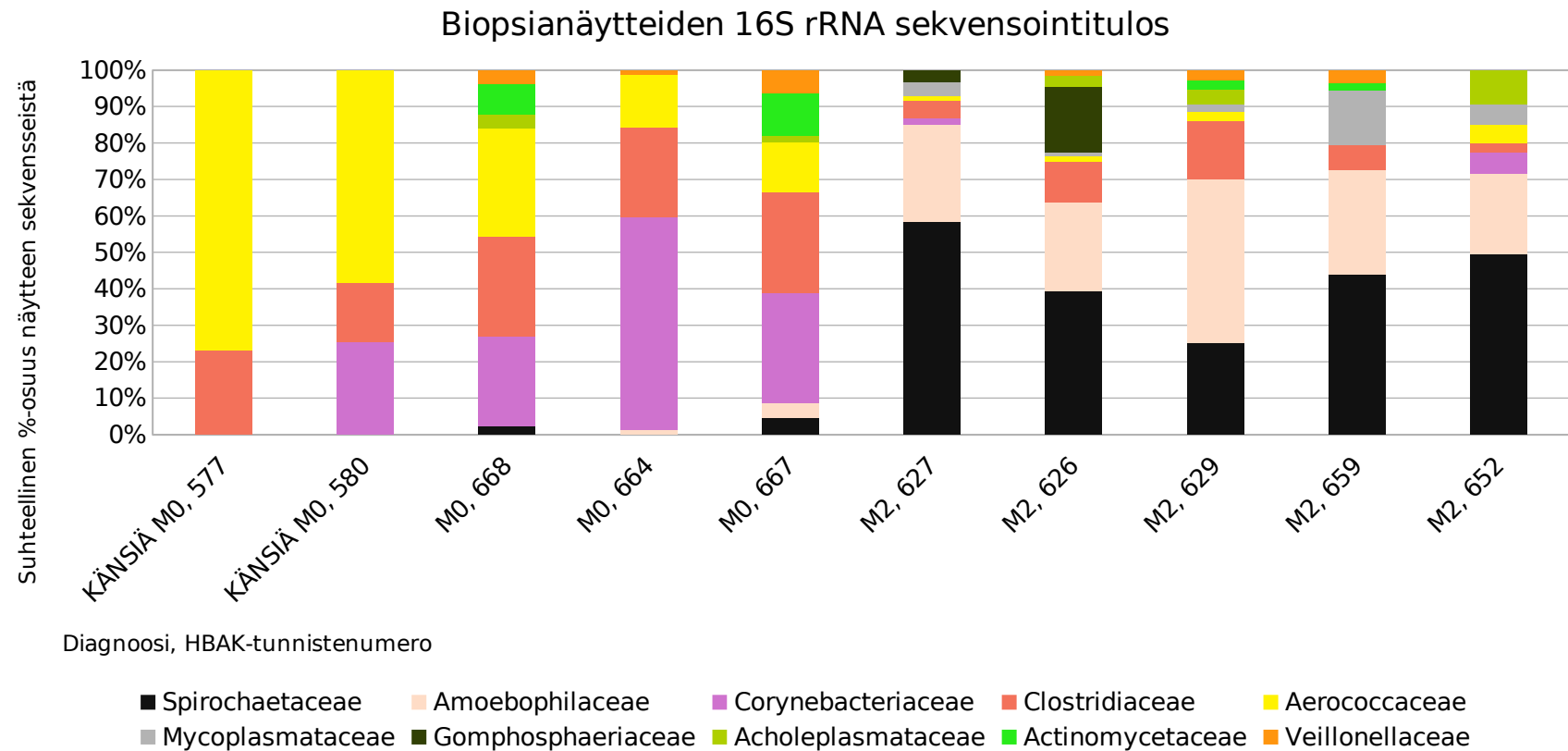
**Kuva 5.** 16S rRNA sekvensointitulos, Chipster-laskenta, DNA-eristys Blood and Tissue -kitillä. 10 yleisintä bakteeriperhettä, jotka ylittävät 500 sekvenssin virhekynnyksen, alukkeet ja indeksointi Kozich ym. 2013 mukaisesti, MiSeq, V2 2x250 bp, ajopäivämäärä oli 19.11.2018.



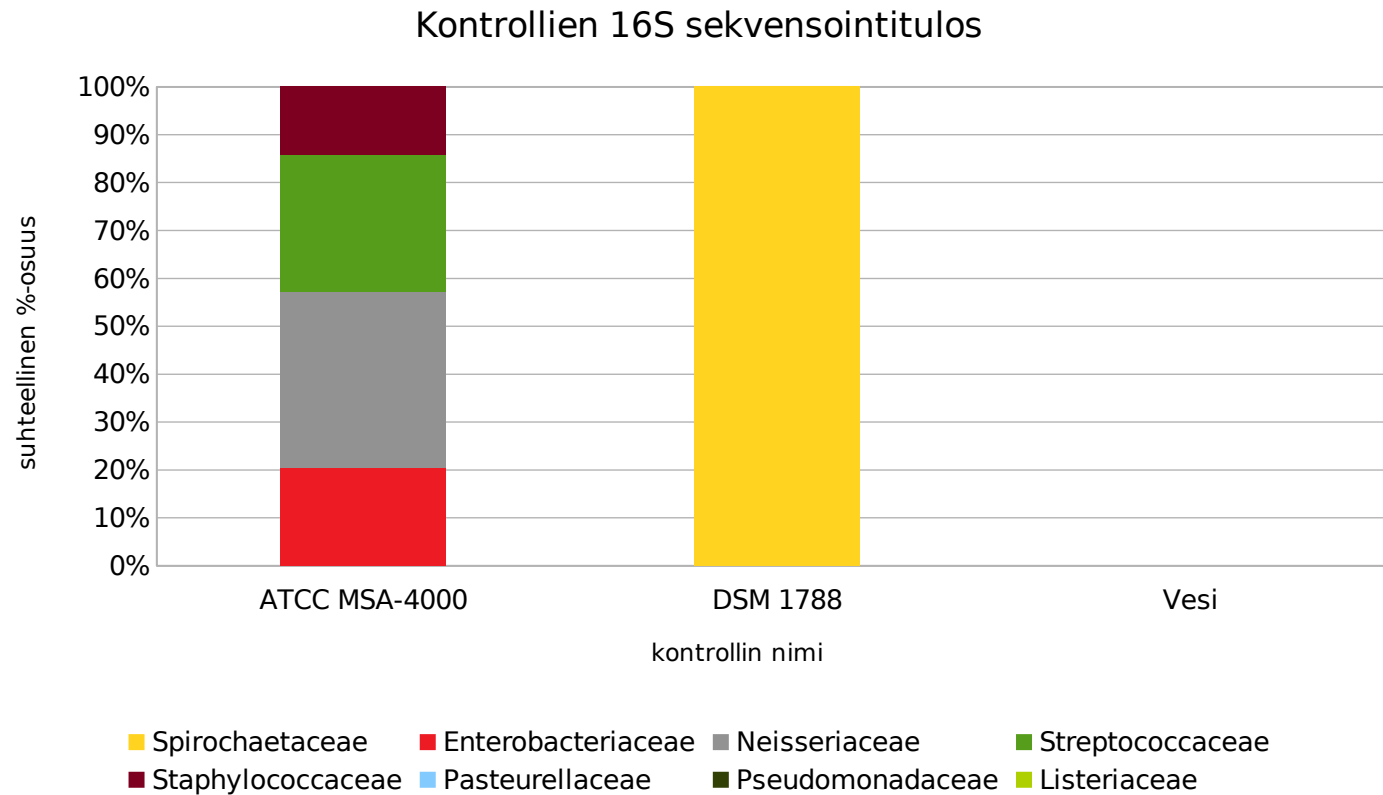
**Kuva 6.** 16S rRNA sekvensointitulokset, Chipster-laskenta, näytteen kuulajahajotus, DNA-eristys Blood and Tissue -kitillä. 10 yleisintä bakteeriperhettä, jotka ylittävät 500 sekvenssin virhekynnyksen, alukkeet ja indeksointi Kozich ym. 2013 mukaisesti, MiSeq, V2 2x250 bp, ajopäivämäärä oli 19.11.2018.



**Kuva 7.** 16S rRNA sekvensointitulos, Chipster-laskenta, DNA-eristysmenetelmä QuickPick SML gDNA -kitillä. 10 yleisintä bakteeriperhettä, jotka ylittävät 500 sekvenssin virhekyynnyksen, alukkeet ja indeksointi Kozich ym. 2013 mukaisesti, MiSeq, V2 2x250 bp, ajopäivämäärä oli 29.5.2019.



**Kuva 8.** 16S rRNA sekvensointitulokset, Illumina 16S Metagenomics-laskenta, DNA-eristysmenetelmä QuickPick SML gDNA -kitillä. 10 yleisintä bakteeriperhettä, jotka ylittävät 500 sekvenssin virhekyynnyksen, alukkeet ja indeksointi Kozich ym. 2013 mukaisesti, MiSeq, V2 2x250 bp, ajopäivämäärä oli 29.5.2019.



**Kuva 9.** Kontrollien 16S rRNA sekvensointitulos, Chipster-laskenta, DNA-eristysmenetelmä QuickPick SML gDNA -kitillä. Yleisimmät bakteeriperheet, jotka ylittävät 500 sekvenssin virhekyynnyksen, alukkeet ja indeksointi Kozich ym. 2013 mukaisesti, MiSeq, V2 2x250 bp, ajopäivämäärä oli 19.11.2018, Chipster-laskenta.

## 6. TULOSTEN TARKASTELU

### 6.1. Näytteenotto, näytteen käsittely ja nautojen hyvinvointi

Hankekarjojen nautojen terveyttä seurattiin tehostetusti biopsianäytteenoton jälkeen eikä yhdelläkään naudoista todettu näytteenotosta aiheutuneita jälkiseuraamuksia. Biopsianäytteenotto ja lääkitseminen oli onnistunut tasalaatuisesti vaihtelevissa kenttäolosuhteissa vuodenajasta huolimatta. Kontaminaatiota tai sekaannuksia näytenumeroinnin suhteen ei ole osoitettu.

Biopsianäytteen kovuus oli erittäin vaihteleva. Monet näytteistä oli otettu usein tulehtuneesta kohdasta sorkka-alueen ihoa ja osin ne olivat hyvin likaisia. Muutamat näytteistä saattoivat sisältää ihon, känsän tai mahdollisen syylän sisään joutuneita kiviä, koska näytettä leikattaessa muutamia skalpellin teriä katkesi.

RNAlater oli eräiden näytteiden osalta epähomogeenista, sitä ei oltu sekoitettu huolellisesti tai taustalla oli muu syy, koska muutaman näyteputken pohjalla havaittiin väritön kidemassa. Biopsianäyte saattoi olla massan sisälle jähmettyneenä.

Näytteen varastoinnissa käytettävä puskuriliuos voi vaikuttaa sekvensoinnista saatavaan lopputuloksen eri bakteerisukujen ja -lajien välisiin suhteisiin eri tavoin (Hallmaier-Wacker ym., 2018). RNAlater-säilytysliuosta kohtaan on myös esitetty kriittisiä näkemyksiä ja mikrobisyhteisöjen tutkimisen osalta näytteen siirtämistä sekunneissa näytteenoton jälkeen -80°C lämpötilaan on suositeltu (Pollock ym., 2018). Tämä vaikuttaa myös näytteenoton työturvallisuuteen ja kustannuksiin.

## 6.2. DNA-eristys

Naudan sorkassa on rustokudosta, joka aiheutti perinteisiä DNA-eristysmenetelmiä käytettäessä vain pienehköt saannot, joten toisen menetelmän käyttöönotto oli ensisijaista. Eristysmenetelmä vaihdettiin pylväspohjaisesta eristysmenetelmästä magneettipohjaiseen eristysmenetelmään ja tämän jälkeen DNA:n saanto parani mutta puhtausparametrit osin alenivat. PCR-tuotteet monistuivat uudella DNA-eristysmenetelmällä pääsääntöisesti tasalaatuisemmin. Eristetyn DNA:n määrä vaihteli runsaasti näytteiden välillä, ja tähän voi myös vaikuttaa nautojen sorkkasairauksien monipuolinen taudinkuva.

Kuulahajotuksella DNA:n määrän keskiarvo oli tällä näyteaineistolla 4,2 ng/μl, normaalilla DNA-eristyksellä keskiarvo oli 13,3 ng/μl. Silputun rustokudoksen ominaisuuksiin ja DNA:n saantoon voitiin kuitenkin vaikuttaa muuttamalla lämpötilaa eristyksen eri työvaiheissa. Kuulahajotus voi osaltaan katkoa DNA-juostetta pieniksi paloiksi ja työhön tulee yksi pipetointivaihe, joten inhimillisen virheen mahdollisuus kasvaa ja tuloksen vertailukelpoisuus voi alentua.

### 6.2.1. Alkuinkuboinnin tarkastelua

DNA-eristykseen sisällytettiin hyvin pitkä solujen hajotusinkubointi (lyysaus). Pohdittiin, tapahtuiko kuuden tunnin +35°C alkuinkuboinnin ja yli yön kestäneen +56°C inkuboinnin aikana bakteerikasvua inkuboinnin keston vuoksi. Näytepala oli ollut näytteenoton jälkeen kylläisessä ammoniumsulfaattiliuoksessa (RNAlater) jääkaapissa ensin vuorokauden, tämän jälkeen sitä säilytettiin -20°C pakkasessa ja se läpikävi pakastus-sulatussyklin, joten biopsianäytteen kudokseen on todennäköisesti jäänyt myös inkuboinnin ajaksi ammoniumsulfaattia, joka on todennäköisesti estänyt bakteerikasvua. Inkuboinnin aikana oli ravistus jatkuvasti päällä,



ja näin mahdollisen ammoniumsulfaatin, Proteinaasi K:n tai RNAasin voidaan olettaa levinneen tasaisemmin Eppendorf-putken sisällä kuin ilman ravistusta. Inkuboinnin aikana lisätty MetaPolyzyme oli eluoitu PBS-liuokseen, joka on puskuriliuos, mutta ei varsinaisesti estä bakteerikasvua. Sekvensointitulokset eivät viittaa suurempaan bakteerikasvuun inkuboinnin aikana lopullisen työvuon osalta, koska tällöin känsäisistä (sorkkasyyläisistä) biopsianäytteistä, joissa DNA:n saanto oli systemaattisen alhainen koko laboratorio-osuuden ajan, tulisi kuitenkin esiintyä runsaammin bakteeriperäistä DNA:ta. Känsäisten (sorkkasyyläisten) näytteiden sekvenssilukumäärät jäivät toistuvasti alhaisemmalle tasolle verraten muihin näytteisiin. Känsäinen tai syyläinen naudan sorkka sisältää pintakerroksiltaan lähinnä vain sarveismassaa. Tällä hetkellä ei ole täysin selvää, ovatko kaikki nautojen sorkassa syyliksi nimitettävät muodostumat virusperäisiä syyliä vai känsiä.

### **6.2.2. Blood and Tissue kitin tarkastelua**

Qiagenin menetelmä käyttää ns. pylväspuhdistusmenetelmää, ja myös tässä tutkielmassa havaittiin agaroosigeeliä uv-valossa tarkasteltaessa pylväspuhdistuksen aiheuttavan DNA:n pilkkoutumista lyhyiksi palasiksi. Pylväspuhdistuksen saanto vaihtelee 50-80% välillä ja tutkimuksia pylväspuhdistuksen korvaamiseksi magneettipuhdistuspohjaisilla menetelmillä on esitetty (Neiman ym., 2012). Tässä tutkielmassa biopsianäytteissä ollut rustokudos saattoi aiheuttaa pylvääseen tukoksia ja alentaa tällä näyteaineistolla DNA:n saantoa. Esimerkiksi pylvään tai eristyksessä käytettävien reagenssien lämmittäminen olisi ollut ongelmallista.

### **6.2.3. Bionobile QuickPick SML gDNA kitin tarkastelua**

Eristysmenetelmä voidaan ohjelmoida esimerkiksi ThermoFisherin KingFisher-pipetointirobottiin. Yhden pesuvaiheen, nestetilavuuksien tarkentamisen, jäähdytysvaiheen lisäämisen, ravistelun lisäämisen sekä eluaatiotilavuuden optimoimisen jälkeen DNA:n saanto oli sekvensointiin kelpoisalla tasolla. DNA oli kuitenkin tällä näytemateriaalilla eri laatuaroja tarkastellen yhä likaista, mutta magneettipartikkelipohjaisella DNA:n eristysmenetelmällä saatiin tämän tutkielman aineistolla suuremmat DNA:n saannot kuin toisella tarkasteltavalla menetelmällä. Lopulta suurin eristetyn DNA:n määrän nousu tapahtui, kun näytettä magneettipartikkeleiden lisäämisen jälkeen inkubointiin +56°C lämpötilassa kevyt ravistelu päällä ja annettiin tämän inkubaation jälkeen kunnolla jäähtyä.

### **6.2.4. Geelieristyksen tarkastelua**

Tässä tutkielmassa havaittiin geelieristysvaiheessa DNA:n saannon systemaattisesti alenevan esimerkiksi 1/5-osaan tai 1/8-osaan Qubit-mittaustulosten perusteella, mutta geelieristyksen automatisointi pipetointirobotin tehtäväksi tai tämän työvaiheen optimointi muulla tavalla voi olla suositeltavaa.

## **6.3. Sekvensointityövuokaavioiden tarkastelua**

1 ng määrä kaksijuosteista DNA:ta, jonka palakoko on 450 bp, sisältää yli  $2 \cdot 10^9$  molekyyliä (kopiota) (Malmström ym., 2005). MiSeqin V2-kittiä käytettäessä yhdessä sekvensointiajossa luetaan miljardeja emäksiä koneellisesti, mutta lukutuloksina V2-kitillä saadaan maksimissaan 15 miljoonaa laatuksiteerit täyttävää sekvenssiä. Menetelmä on ollut käytössä jo vuosia ja synteisiin perustuvan sekvensoinnin kemia sekä muut ominaisuudet tunnetaan hyvin.

Illuminan työvuon sekvenssilukumäärät olivat alhaiset tässä tutkielmassa, ja syy tähän voi myös olla esimerkiksi lopullisen latausliuoksen kvantifioinnissa, tosin tämä olisi pitänyt tapahtua kahdesti. Illumina 2 -ajon osalta indeksi 1-tunnistamisen laatuarvot olivat hyvin alhaiset. Myös PCR-reaktion herkkyys templaatti-DNA:n konsentraation vaihteluille on riskinä Illuminan protokollassa, koska työvaiheiden uusinta on työlästä ja alttius virheille kasvaa.

Kozichin menetelmän indeksijärjestelmä toimi tässä tutkielmassa paremmin kuin Illuminan kahden PCR-reaktion menetelmä, koska tulokseksi saadut, tunnistamisen läpikäyneiden sekvenssien lukumäärät olivat selvästi suuremmat ja näytteet monistuiivat tasalaatuisemmin. Myös mahdollisen kimeran tai yleisesti epäspesifisen sitoutumisen todennäköisyys on alempi (Madigan ym., 2019). Sekvensoinnista saatavien sekvenssien lukumäärä vaikuttaa myös tulosten käyttökelpoisuuteen, koska bakteerien diversiteettiä tarkasteltaessa ( $\alpha$ -,  $\beta$ - tai  $\gamma$ ) tarvitaan runsas määrä sekvenssejä näytettä kohden luotettavan tuloksen saamiseksi (Pollock ym., 2018). Tämän tutkielman sekvensoinneissa saatiin eniten sekvenssejä kaupallisista kontrollinäytteistä, joissa DNA:n konsentraatio PCR-reaktion jälkeen oli korkein ja puhtausparametrit olivat hyvällä tasolla.

### **6.3.1. Sekvensoinnit V2 ja V3-tuotteilla**

Tutkielman esikokeissa tehtiin sekvensointi Illuminan V3-kitillä, jolla saadaan 2x300 emäsparin pituisia sekvenssejä näytteistä. V3-kitin sekvensoinneista saadun sekvenssin laatu oli noin 250 emäksen jälkeen alhainen ja jäi alle Phred Score 20 -arvon. V3-kitillä sekvenssi 2 (read2) oli MultiQC-raporttien perusteella huonolaatuinen. Chipster-ohjelmisto jo analysoinnin alkuvaiheessa hylkäsi suurimman osan näistä

huonolaatuisista sekvensseistä ja tämän vuoksi siirryttiin käyttämään sekvensoinnissa Illuminan V2-kittiä.

### **6.3.2. Alukkeet ja PCR-reaktiot**

DNA-eristysmenetelmän vaihtamisen jälkeen eivät PCR-reaktiot aluksi toimineet. PCR:n toimivuuteen vahvasti vaikuttavaa DNA:n määrää jouduttiin kokeellisesti tarkastelemaan muutamaa otteeseen ja DNA:n laimennos PCR-reaktiota varten todettiin toimivaksi menetelmäksi. Tämän perusteella konsentroituneita DNA-templaatteja päätettiin laimentaa korkeintaan 30 ng/μl-pitoisuuteen.

Tämän tutkielman osalta osa 29.5.2019 sekvensoinnin tuotteista oli laimennettu, osa ei. Syy tähän on aikataulullinen, mutta toisaalta esimerkiksi DNA-eristyksen ja koko sekvensointityövuon vaihtamista toiseen ei voitu ennustaa etukäteen.

PCR-reaktioiden toimivuuteen voivat myös vaikuttaa näytteessä esiintyvät inhibiittorit. Näitä ovat orgaanisten aineiden osalta uloste, maaperän humus tai turve sekä polysakkaridit, epäorgaanisista aineista mm. kalsium-ionit (Pollock ym., 2018).

Kaikki arkeonit ja bakteerit täydellisesti ja luotettavasti tunnistavia yleisalukkeita ei toistaiseksi ole esitetty, ja valittavien alukkeiden sitoutumisalue vaikuttaa myös tulokseen (Pollock ym., 2018). ATCC MSA-4000 -kontrollin avulla saatiin tietoa sekvensoinnin bakteereista, jotka ylitai aliedustuivat. Kozichin menetelmän alukkeilla saatiin osoitettua melko runsaasti eri bakteeritaksoneita, mutta alukkeet ovat melko pitkät, joka voi vaikuttaa PCR-reaktion toimivuuteen.

#### **6.4. Kirjaston valmistus ja näytteen lataus**

MiSeqin 16S rRNA-sekvensointimenetelmän näytemäärä on yleensä 50-384 kpl, joten mahdollisiin pipetointivirheisiin tulisi kiinnittää huomiota jo aikatauluja sekä työvaiheita suunniteltaessa. Edelleen näytteen latauksessa jo pienetkin konsentraatioerot voivat vaikuttaa MiSeqin flowcell-osassa muodostuvaan klusterointitiheyteen ja näytteistä saataviin sekvenssilukumääriin (Kozich ym., 2013). Tutkielman havaintojen perusteella latauskonsentraatiota kannattaa tarkastella sekä tarvittaessa muuttaa ajokohtaisesti, mikäli sekvensoitavan DNA:n laatu vaihtelee.

#### **6.5. Qubit-, DeNovix ja Bioanalyzer-mittalaitteet**

Bioanalyzer-mittalaitteen osalta havaittiin tulosten graafisessa esityksessä eräs ohjelmistovirhe, joka ei ollut laitteen automaattisten päivitysten myötä korjaantunut: tulossivulla oli tuhatkertainen ero konsentraatioissa pikomoolien muututtua nanomooleiksi. Bioanalyzer-mittalaite on hyvin herkkä mittalaite ja laitteen reagenssienkin suhteen kannattaa suorittaa säännöllistä tarkastelua. Bioanalyzer-mittalaite osoittautui käyttökelpoisimmaksi sekvensointiin laitettavan, lopullisen yhdistelmänäytteen (kirjaston) mittaamisessa ja valinnassa, mutta myös tällöin rinnakkaismittausta Qubitilla suositellaan.

Mittausepävarmuuksista on huomattava myös spektrofotometriaan perustuvan DeNovix-mittalaitteen virhe ja mitattava matriksi. Työvuon kontrollien tulee olla kattavat ja sisältää sekä vesi- että eri reagenssikontrolleja.

Fluoresenssiin perustuva, kaksijuosteista DNA:ta mittaava Qubit-mittaus on osoittautui tarkaksi menetelmäksi tässä tutkielmassa. Valmistajan ohjeita reagenssien lämpötiloista tulee noudattaa.

## 6.6. Bioinformatiikkaosuuden tarkastelua

Eri sekvensointityövuokaavioilla sekvenssien päällekkäisen, sekvensaattorin lukeman alueen pituus (engl. overlap) voi vaihdella, ja konsensussekvenssiä muodostettaessa mahdollisimman suuri päällekkäinen osuus on suositeltavaa ja tuottaa parempia tuloksia (Kozich ym., 2013). Tämä päällekkäisyys on Kozichin menetelmässä Illuminan menetelmää suurempi.

Chipster-ohjelmistopaketti toimi tämän tutkielman tarkasteluissa luotettavasti. Chipsterillä saavutetaan *Treponema*-bakteerien osalta SILVA-tietokantaa vasten linjatessa tällä hetkellä sukutason varmuus, mutta yksittäisten sekvenssien tarkastelu BLAST-tietokantaa vasten on edelleen mahdollista. Kuitenkin komentorivipohjaisten ohjelmien käyttö BLAST-tietokannan tarkastelussa on nopeampaa kuin nykyistä Chipster-versiota käytettäessä.

Chipsterillä myös laskettiin 29.5.2019 sekvensoinnin tulokset uudelleen "Max. number of ambiguous bases = 0" -arvolla, jolloin tulos oli kymmenen yleisimmän bakteeritaksonin osalta pääosin sama paitsi *Bacteroidia* unclassified ja *Bacteria* unclassified -annotaatiot olivat keskenään eri järjestyksessä. Yksittäisen näytteen osalta eroavaisuuksia voi tällä tarkastelutapojen erolla tulla. Sekvenssilukumäärät olivat vain hieman alhaisempia. Emästen täsmällinen tunnistaminen on edellytys täsmällisille tuloksille, joten arvo 0 on jatkossakin suositeltavampi. Tulosten uudelleenlaskenta linjauksen esiklusteroinnin asetuksella "Number of differences allowed = 0" ja edelleen tuloksen vertaaminen lajitasolla arvolla 1 laskettuun voi myös olla mahdollinen jatkotoimenpide.

Eri bakteeritietokantojen osalta Robert C. Edgarin mukaan monet sekvenssitiedot yhdistettynä taksonomiseen tietoon (annotaatiot) ovat

tällä hetkellä usein ennusteita (engl. predicted), eivät todellisia, varmistettuja sekvenssejä. SILVA- ja Greengenes-tietokantoja keskenään verratessaan Edgar oli löytänyt 249490 identtistä sekvenssiä, joiden annotaatiot olivat ristiriitaiset. Edgar päätyi arvioon, että vuonna 2018 SILVA:n versio v.128:n ja Greengenes v.13.5 osalta noin 17% annotaatioista sisälsi virheitä ja molempien tietokantojen alempi virheraja olisi vähintään 34%. SILVAN osalta tämän tutkielman osalta käytettiin versiota v.132 (Edgar, R. 2018).

BLAST-tietokanta ei ole kuratoitu ja sinne ladataan myös koko ajan sekvensointitietoa, jonka laatukontrollitiedot ovat puutteellisia.

Tässä tutkielmassa myös havaittiin, että vaikka sekvensointiajon laatuarvo Phred score alitti arvon 20 ja Chipster-ohjelmapaketti hylkää huonolaatuiset tulokset, niin Illuminan 16S Metagenomics-työkalulla kuitenkin saatiin annotaatioita laatuarvot alittaneista sekvensseistä, eikä syy tähän ole täysin selvillä.

## **6.7. Tulosten esitys ja rarefaktio**

Tulosten esityksen osalta rarefaktiota, tulosten arpomista, on perusteltu argumentein, että näin voidaan näyttöä, joiden sekvenssilukumäärät vaihtelevat runsaastikin, verrata keskenään. Väitettä voidaan pitää osittain totena teoriatasolla, mutta on kuitenkin myös esitetty, että rarefaktiossa on ongelmia ja siitä tulisi luopua (McMurdie & Holmes, 2014). Virheen arvioinnissa koko työvuon analysointi kokonaisdifferentiaalien avulla sekä koko ajon sekvenssilukumäärien keskihajontojen tarkastelu tai myös niiden ilmoittaminen saattaisi olla eräs toimintamalli.

## 6.8. Bakteriologisten tulosten tarkastelu

Sekvensoinneista laskettuja tuloksia tarkasteltiin ja vertailtiin kirjallisuuteen. *Treponema*-suvun bakteereita saatiin osoitettua eri sekvensoinneissa vaihtelevia määriä, osasta näytteitä runsaasti. Jatkotutkimuksena tuloksen varmentamista lajitasolle esimerkiksi kolmannesta tietokannasta suositeltaisiin.

### 6.8.1. Vertailu kuulahajotuksen sekä normaalin DNA-eristyksen välillä

DNA-eristysmenetelmänä 19.11.2018 sekvensoinnissa oli Blood and Tissue -kitti, Chipster-laskenta.

Normaalilla DNA-eristyksellä saatiin tällä aineistolla myös suurempi sekvenssilukumäärä (keskiarvona 65800 sekvenssiä/näyte) kuin kuulahajotusta käyttämällä (47600 sekvenssiä/näyte). Sekvensointitulosten tarkastelun perusteella kuulahajotuksella saatiin hieman runsaammin virhekynnyksen ylittäviä eri bakteeritaksoneita kuin normaalilla DNA-eristysmenetelmällä. Normaalilla DNA-eristyksellä yhdestä näytteestä saatiin sekvenssejä vain niukasti virhekynnyksen, kuulakäsittelyllä samasta näytteestä saatiin taas runsaammin sekvenssejä. Käksi tai syyliä sisältävien näytteiden osalta sekvenssilukumäärät ja eri bakteeritaksonien määrät jäivät alhaisiksi molemmilla DNA-eristysmenetelmillä.

*Spirochaetaceae*-perheen bakteereita osoitettiin kuulakäsittelyistä näytteistä kaikista naudoista, normaalilla DNA-eristyksellä vain seitsemästä. *Spirochaetaceae*-perheen bakteerien määrät vaihtelivat noin 5-50%:n välillä näytekohtaisesti ja Chipsterin tuloksen mukaan ne olivat *Treponema\_2* ja *Spirochaetaceae* unclassified-taksoneita.



*Clostridiales*-lahkon perhe XI ja *Clostridiales* unclassified -taksoneita havaittiin eri näytteistä runsaasti molemmilla eristysmenetelmillä, *Actinomycetaceae*-perhettä osoitettiin pieniä määriä kuudesta näytteestä. *Corynebacteriaceae*-perhettä osoitettiin seitsemästä näytteestä vaihtelevia määriä *Mycoplasmataceae*-perhettä osoitettiin viidestä näytteestä, joiden kaikkien diagnoosi oli M4 tai M4.1. *Aerococcaceae*-perhettä osoitettiin seitsemästä näytteestä. *Porphyromonadaceae*-perhettä osoitettiin kuudesta eri M4 tai M4.1. -näytteestä. *Ruminococcaceae*-perhettä osoitettiin seitsemästä näytteestä, *Carnobacteriaceae*-perhettä kuudesta eri näytteestä.

#### **6.8.2. Vertailu Chipster-laskennan ja 16S Metagenomics-laskennan välillä**

DNA-eristysmenetelmänä 29.5.2019 sekvensoinnissa oli Bionobile QuickPick SML gDNA kit.

*Spirochaetaceae*-bakteeriperhettä osoitettiin Chipster-laskennan osalta (Kuva 7) pieniä määriä kahdesta M0- ja runsaasti (~20-50%) kaikista M2-naudoista, Illuminan laskenta (Kuva 8) antoi samanlaiset tulokset. Chipsterin tuloksen mukaan nämä taksonit olivat miltei kokonaan *Treponema\_2* ja *Spirochaetaceae* unclassified-taksoneita.

Molempien laskentojen osalta *Corynebacteriaceae*-perheen bakteereita havaittiin eniten M0-naudoilla. *Clostridiaceae*-perhettä havaittiin molemmilla laskennoilla vaihtelevia määriä eri naudoilta. *Mycoplasmataceae*-perhettä osoitettiin kaikilla M2-diagnoosin saaneilla naudoilla mutta ei M0-naudoilla, tulos oli sama molemmilla laskentatavoilla. *Aerococcaceae*-perheen taksoneita havaittiin molemmilla laskentatavoilla hieman runsaammin M0-naudoilla. *Actinomycetaceae*-bakteeriperhettä havaittiin sekä M0- että M2-naudoilla, nämä tunnistuivat

tällä tavalla vain Illuminan laskennalla. Illuminan laskennalla kaikilla M2-naudoilla ja yhdellä M0-naudalla osoitettiin *Amoebophilaceae*-bakteeriperheen taksoneita ja *Amoebophilaceae*-perheen osalta laskenta on tarkistettava edelleen esimerkiksi BLAST- ja RDPII-hauilla, eri tietokannat antavat tälle taksonille erilaisen vastauksen ja ristiriita on samankaltainen kuin kahdessa nautojen sorkkaterveyttä käsittelevässä tutkimuksessa, Zinicola ym., 2015 ja Nielsen ym., 2016.

Fusobakteerit puuttuivat molemmilla laskentatavoilla kymmenen yleisimmän bakteeritaksonin joukosta. Tämä oli osin yllättävää ja asia saattaa vaatia laajemman näyteaineiston kuin tässä tutkielmassa. Eräs selitys tähän voi olla teoria, että ID ja DD ovat kaksi eri tautia, koska fusobakteereita ID:ssä on osoitettu viljelymenetelmin (Kontturi ym., 2019). Lisäksi taudeilla voi olla maantieteellisiä eroja (Madigan ym., 2019). Tämän tutkielman naudat oli diagnosoinut DD:ksi kokenut suomalainen nautaeläinlääkäri.

Kaiken kaikkiaan eri tietokannat antoivat erilaisia annotaatioita jo perhetasolla, joten tuloksen tarkastelu kahteen avoimeen tietokantaan lajitasolle saakka sekä tuloksen selvittely eri R- ja keinoälysovelluksilla olisivat loogisia seuraavia jatkotoimenpiteitä.

### **6.8.3. Kontrollit**

29.5.2019 sekvensoinnissa ATCC MSA-4000-kontrollin osalta joitakin bakteerilajeja saatiin alle tämän tutkielman virhekynnyksen, ja tarkemman virhekynnyksen asettaminen voi vaatia lisäselvitystä. DSM 1788-kontrollista saatiin Chipster-laskennalla 529 734 *Spirochaetaceae*-perheen sekvenssiä, Illuminan laskennalla saatiin 552 815 *Spirochaetaceae*- sekvenssiä.

Erot voivat johtua myös Kozichin menetelmän alukkeista, mutta DSM 1788 -kontrollituotteen laimentaminen jatkossa 1:10 tai 1:100 voi olla tarpeen.

Tässä työssä havaittiin kontrollien aiheuttavan kustannusten nousua mutta nopeuttavan sekä varmentavan koko sekvensointitulosten tarkastelua työvuon eri vaiheissa kuten kirjallisuudessa (McGovern et al., 2018).

### **6.9. 16S rRNA-geenin sekvensoinnin käyttökelpoisuus, jatkotutkimukset**

16S rRNA-geenin alueen sekvensointi on vertailukelpoisia tuloksia tuottanut menetelmä (Madigan ym., 2019; Pollock ym., 2018). Kuitenkin nykyteknologialla sekvensoinnissa on virhelähteitä, joita on myös tutkittu, ja emäsjärjestykseen perustuvasta indeksoinnista on osoitettu indeksin väärintunnistumisesta johtuvaa virhettä 2.1-2,6 % (Schnell ym., 2015) tai huomattavasti runsaammin (Esling ym., 2015). Kozichin mukaan virhe on vähintään 0,1% (Kozich ym., 2013). Näin ollen virhekyynnyksen asettaminen sekä ja näytteiden tekniset toistot olisivat suositeltavia. Tuloksen rarefaktiossa voidaan algoritmin asetuksista riippuen valita sattumalta sekvenssi, joka edustaa hyvin huonosti näytteen mikrobiomia, joten esimerkiksi kymmeneen yleisimpään bakteerisukuun ja -lajiin perustuva tarkastelu voi edustaa paremmin näytettä.

PCR-pohjaisen 16S rRNA-sekvensoinnin kvantitatiivisuuteen vaikuttaa moni asia työvuossa, valituista alukkeista alkaen. Useimmissa sekvensaattoreissa tapahtuu myös uusi PCR-reaktio, mutta toisaalta kvantitatiivisuus on yleinen immunologian sekä bakteriologian ongelma: miten rajataan tai määritellään esimerkiksi tulehdusalue tarkasti naudan sorkassa?

Mikäli sekvensointituloksissa ei esiinny jotain bakteeritaksonia, se ei toistaiseksi todista sen poissaoloa, sillä näytteenotossa, laboratoriotyössä, PCR-reaktioissa sekä sekvensoinneissa on paljon muuttuvia tekijöitä. Näytemateriaali tai sitä ympäröivä ympäristö saattaa sisältää PCR-reaktioita inhiboivia aineita. Mikäli näytteestä saadaan virhekyynnyksen ylittävä määrä sekvenssejä ja kahdesta eri avoimesta tietokannasta sama taksonitulos, on se vahva osoitus bakteerin olemassaolosta näytteessä, mikäli koko ketju näytteenotosta sekvenssaattorin lataamiseen on kunnossa ja kontaminaatioita ei ole. Bakteerin määrällinen arviointi voi olla vaikeaa, koska näytteenotossa ja eri sekvensointilaitteistossa on myös välivaiheita, jotka Illuminan teknologiassa tunnetaan hyvin.

Tässä työssä valitun Kozich ym. 2013 työvuon alukkeet saattavat monistaa treponemabakteereita osin huonommin kuin Klitgaardin tutkimuksen lyhyet alukkeet, koska myös alukkeiden pituus voi vaikuttaa PCR-reaktion onnistumiseen. Jatkotutkimuksina ovat ensisijaisesti *Treponema*-bakteereiden osalta tuloksen varmistaminen BLAST- ja ehkä myös RDPII -tietokantaa vasten sekä uusimpien *Treponema*-bakteereiden taksonomisten määrittelyjen mukainen tulosten tarkastelu. QIIME2-, VSEARCH ja V-Xtractor-ohjelmistojen kokeilu sekvensointituloksen analysoinnissa olisi myös kokeilemisen arvoinen. Eri kaupallisilla keinoälysovelluksilla sekä R-ohjelmalla voidaan myös tulosten tarkastelua sekä visualisointia saada nopeammin muovattavalle ja tehokkaammalle tasolle.

Virheentarkastelun osalta kokonaisdifferentiaalien muodostaminen koko työvuon ja sekvensointituloksien virheelle voisi tarjota näkökulman tuloksen luotettavuuteen.

Yhdysvaltalaisessa eläinlääketieteen oppikirjassa todetaan "The odor of foot rot is strong and characteristic", eli jo sorkan ajotulehduksen (ID) sisällä on todennäköisesti useita eri, jopa ehkä tarkalla hajuaistilla havaittavia eroja bakteriologisissa taudinaiheuttajissa (Peek & Divers, 2018). Olisiko tämä sama mahdollinen myös DD:n osalta? Voivatko ID ja DD olla yleensä eri, mutta joskus myös saman tai samojen taudinaiheuttajien aiheuttamia? Voidaan toki kysyä, onko tämän tutkielman aineiston ja tuloksen perusteella saatu viitteitä DD:n ja muiden sorkkasairauksien jakautuvan Suomessa bakteriologisesti useampaan eri tautiryhmään kuin aikaisemmat tutkimukset ovat osoittaneet.

## 7. JOHTOPÄÄTÖKSET

Tutkielman tavoitteiden osalta

- i) Menetelmäpystytys onnistui Kozichin työvuolla ja lopulliseen 29.5.2019 sekvensointiin olisi kannattanut ottaa kaikki näytteet kolmena toistona. Tärkeimmät kontrollit tulisi tehdä myös kolmena toistona.
- ii) Tässä tutkielmassa havaittiin eläinlääkärin terveiksi luokittelemien nautojen (M0) sorkkien ihoalueen mikrobiomin eroavan M2-diagnoosin nautojen sorkkien mikrobiomista. M2-diagnoosin saaneilla naudoilla havaittiin eniten *Spirochaetaceae*-perheen bakteereita, joita tämän tutkielman M0-naudoilla osoitettiin hyvin vähän tai ei lainkaan. Jatkotutkimuksena nyt saatujen sekvensointitulosten vertaaminen BLAST- ja RDP II- tietokantaa vastaan olisi suositeltavaa. *Amoebophilaceae*-perheen taksonien osalta oli ristiriitaisuuksia. *Mycoplasmataceae*-perheen bakteereita osoitettiin enemmän M2-diagnoosin naudoilla, ja tähän bakteeriperheeseen kuuluu esimerkiksi tulehduksissa yleisesti esiintyviä bakteereita. *Actinomycetaceae*-perheen bakteereiden esiintyvyys ja navetan kuivikelaadun tarkastelu voisi myös olla lisäselvitysten kohteena.

Eri *Porphyromonas*- ja *Fusobacterium*-sukujen bakteerien sekä *Corynebacteriaceae*-bakteeriperheen esiintyvyyttä kannattaa myös lisäselvittää suuremmalla aineistolla, sillä tutkielman tulos ei näidenkään taksonien osalta pääosin vastaa viljelymenetelmin saatuja tuloksia.

- iii) Lähes kaikki sairailta eläimiltä osoitetut *Spirochaetaceae*-perheen bakteerit olivat eri *Treponema*-bakteereita, niitä osoitettiin 20-50% näytteen bakteerien kokonaismäärästä.

Lisäksi havaittiin, että biopsianäytteenotto kenttäolosuhteissa on kallista ja sisältää keskimääräistä suuremman riskin naudän sorkan infektoitumisesta näytteenoton yhteydessä, mutta tällä menetelmällä saatiin todennäköisesti osoitettua myös ihon sisällä tai ihonalaiskudoksessa sijaitsevia bakteereita. Varsinkin diagnostisessa tarkastelussa sekvensointituloksen saamiseen käytettävä aika tai rahamäärä voi olla rajallinen.

Illuminan omalla, kahden PCR-reaktion 16S-työvuolla tässä tutkielmassa laadukkaiden sekvenssien lukumäärä jäi hyvin alhaiseksi ja asetettuun tavoitteeseen ei päästy. Kozichin työvuolla sekvenssien lukumäärä oli lähellä tavoitetta, laatuarvot olivat molemmissa menetelmän sekvenssoinneissa paremmat ja sekvensseissä on enemmän päällekkäisyyttä (overlap).

Bioinformatiikan osalta avoimen lähdekoodin ohjelmilla tapahtuva, saatujen sekvensointitulosten vertaaminen kahteen erilliseen, avoimeen tietokantaan tuottaa tieteellisen tiedon kriteerejä täyttäviä tuloksia. Tuloksia on edelleen syytä tarkastella avoimesti sekä tieteellisen kriittisesti. Avoimen lähdekoodin ohjelmistoilla esimerkiksi väärän positiivisen tai virheellisen negatiivisen tuloksen tilanteessa koko bioinformatiikkatyövuon tarkastelu on mahdollista.

Kaiken kaikkiaan suomalaisten nautojen eläintautitilanne on ollut pitkään kansainvälisesti arvioiden hyvällä tasolla, lääkkeiden käyttö on vähäistä ja jokaisella naudalla on huolenpitoa varmistava persoonallinen nimi, joten

niiden sorkkaterveyden parantamiseen tähtääviä toimenpiteitä on hyvinkin mahdollista toteuttaa.



## **Huomiot ja kiitokset**

Työssä havaitut virheet ovat yksin kirjoittajan tekemiä.

Erityiskiitokset kuuluvat yksikönjohtaja, tutkimusprofessori Sinikka Pelkoselle, FT, tutkija Mikko Lehtoselle ja myös tämän työn ohjaajalle, MMT (mikrobiologia), erikoistutkija Taru Lienemannille Ruokavirastossa, joiden kaikkien panos tämän tutkielman onnistumiseen sekä ongelmanratkaisuihin oli lopulta huomattava. Kiitosten osalta huomautettakoon, että menetelmänpystytystyössä saattaa ajoin ilmetä ongelmallisia tilanteita, mutta esimiesten suoraviivaisen vahva tuki tämän tutkielman valmistumiseksi edesauttoi runsaasti työn valmistumista.

Tutkimusyhteisössä koko työyhteisön ja myös naapuriosastojen tutkimusmyönteinen toimintakulttuuri vaikutti tässä tapauksessa hyvin positiivisesti työn valmistumiseen. Kiitos siis mahdollisuudesta tehdä tämä työ Ruokaviraston organisaatiossa, jonka BAPA- ja MIBO-osaston välle osoitan kiitokset, pedanteista laitoshuoltajista tiskareihin sekä esimiehiin asti, koska ilman pitkäjänteisyyttä arvostavaa, kokenutta ja ammattitaitoisesti työnsä tekevää henkilöstöä tämä työ ei olisi onnistunut tai valmistunut. Laboratorin osalta varsinkin Kirsi-Maria Eklundille sekä Riikka Mäkiselle suuri kiitos avustanne! Tutkielman valmistumista avustivat ja edistivät myös Ruokaviraston Helsingin eläinpatologian jaostolle sekä Helsingin yliopiston Saaren toimipiste, joille myös osoitetaan kiitos.

Maa- ja metsätalousministeriön koordinoimaan hankkeeseen osallistuivat myös Valio Oy, Eläinten terveys ETT ry, Helsingin yliopiston eläinlääketieteellinen tiedekunta ja Ruokavirasto, jotka kaikki ansaitsevan kiitoksen perustutkimuksen rahoittamisesta. Kiitos myös kaikille Sorkka

DD -hankkeen henkilöille monista keskusteluista ja oppimissävytteisistä tilanteista, kuuntelukyvystänne sekä varsinkin kärsivällisyydestänne!

CSC:n tietojärjestelmien sekä kaikkien avoimen lähdekoodin ohjelmien toimivuus oli tutkielman suorituksen aikana luotettavalla ja hyvällä tasolla. Kiitoksen ansaitsevat CSC:n Kimmo Mattila ja eritoten FT Eija Korpelainen.

Kiitos vanhemmilleni ja koko muulle lähipiirilleni, myös edesmenneille sedälleni ja kummisedälleni kaikesta tuesta, jota vuosien varrella olen saanut.

I would like to thank my (mostly younger) fellow students for many usefull conversations and lunch-discussions (scientific & off-topic), which usually created positive effects on my thinking, although sometimes this also meant a lot of extra work in the lab or in front of the computer screen! Suurkiitos myös koko Savolaisen osakunnan laulajien porukalle tämän työn valmistumisen aikana, jolloin viikoittainen annos kuorolaulua vei ajatukset sujuvasti pois arjesta.

Useat opettajat ovat myös tehneet työnsä niin, että intoa luonnontieteisiin saattoi vielä tässä vaiheessa olla jäljellä. Fyysikko Matti Aaltosen tinkimättömät, matemaatikko Kyösti Tarvaisen hiljaisuutta vaativat ja dosentti Pekka Heinon hyvin innostavat opetusmenetelmät ovat painuneet syvälle mieleen, lisäksi Vappu-Liisa Hakalehdon ja Lotta-Maija Kuusojan opetuksen persoonallisuus, vaikuttavuus ja energisyys tulee myös kiitoksin mainita.

”Olisi ymmärrettävä, että fanaattinen voitontavoittelu ja oman arvon ansaitseminen suorituksilla on yhä paheneva sairaus suomalaisessa yhteiskunnassa. Ei riitä, että on se mikä on: tavallinen ihminen, joka löytää elämästään paljon hyvää. Parasta ovat hyvät ihmissuhteet, kaverit ja eri-ikäiset ystävät.” -Kari Uusikylä

"No man should escape our universities without knowing how little he knows."

-J. Robert Oppenheimer

## LÄHDELUETTELO

- Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). *Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton*. *Aquatic Microbial Ecology* (Vol. 75). Retrieved from [www.int-res.com/articles/suppl/a075p129\\_supp.xlsx](http://www.int-res.com/articles/suppl/a075p129_supp.xlsx)
- Baron, E. J., & Murray, P. R. (1999). *Manual of clinical microbiology* (7. ed). Washington, D.C.: ASM Press.
- Berry, S. L., Read, D. H., Famula, T. R., Mongini, A., & Döpfer, D. (2012). Long-term observations on the dynamics of bovine digital dermatitis lesions on a California dairy after topical treatment with lincomycin HCl. *Veterinary Journal*, 193(3), 654–658. <https://doi.org/10.1016/j.tvjl.2012.06.048>
- Bos, K. I., Herbig, A., Sahl, J., Waglechner, N., Fourment, M., Forrest, S. A., ... Poinar, H. N. (2016). Eighteenth century *Yersinia pestis* genomes reveal the long-term persistence of an historical plague focus. *ELife*, 5, 1–11. <https://doi.org/10.7554/elifelife.12994>
- Clegg, S. R., Carter, S. D., Birtles, R. J., Brown, J. M., Anthony Hart, C., & Evans, N. J. (2016). Multilocus sequence typing of pathogenic treponemes isolated from cloven-hoofed animals and comparison to treponemes isolated from humans. *Applied and Environmental Microbiology*, 82(15), 4523–4536. <https://doi.org/10.1128/AEM.00025-16>
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., ... Tiedje, J. M. (2014). Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, 42(D1), 633–642. <https://doi.org/10.1093/nar/gkt1244>
- Demirkan, I., Erdoğan, M., Demirkan, A. Ç., Bozkurt, F., Altındış, M., Navruz, F. Z., & Köse, Z. (2018). Isolation and identification of *Treponema pedis* and *Treponema phagedenis*-like organisms from bovine digital dermatitis lesions found in dairy cattle in Turkey. *Journal of Dairy Science*, 101(11), 10317–10326. <https://doi.org/10.3168/jds.2017-14227>

- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., ... Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, 72(7), 5069–5072.  
<https://doi.org/10.1128/AEM.03006-05>
- Döpfer, D., Koopmans, A., Meijer, F. ., Szakáll, I., Schukken, Y. H., Klee, W., ... ter Huurne, A. A. H. M. (1997). Döpfer ym. 1997. *Veterinary Record*, 140, 620–623. Retrieved from 10.1136/vr.140.24.620
- Dupuy, V., Manso-Silván, L., Barbe, V., Thebault, P., Dordet-Frisoni, E., Citti, C., ... Thiaucourt, F. (2012). Evolutionary History of Contagious Bovine Pleuropneumonia Using Next Generation Sequencing of *Mycoplasma mycoides* Subsp. *mycoides* "Small Colony." *PLoS ONE*, 7(10).  
<https://doi.org/10.1371/journal.pone.0046821>
- Edgar, R. (2018). Taxonomy annotation and guide tree errors in 16S rRNA databases. *PeerJ*, 2018(6). <https://doi.org/10.7717/peerj.5030>
- Edgar, R. C. (2018). Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ*, 2018(4).  
<https://doi.org/10.7717/peerj.4652>
- Esling, P., Lejzerowicz, F., & Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research*, 43(5), 2513–2524. <https://doi.org/10.1093/nar/gkv107>
- Evans, N. J., Murray, R. D., & Carter, S. D. (2016, May 1). Bovine digital dermatitis: Current concepts from laboratory to farm. *Veterinary Journal*. Bailliere Tindall Ltd. <https://doi.org/10.1016/j.tvjl.2015.10.028>
- Frössling, J., Rosander, A., Björkman, C., Näslund, K., & Pringle, M. (2018). Detection of *Treponema phagedenis*-like antibodies in serum and bulk milk from cows with and without digital dermatitis. *Journal of Veterinary Diagnostic Investigation : Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, 30(1), 86–92.  
<https://doi.org/10.1177/1040638717733778>

- Graves, S., & Downes, J. (1981). Experimental infection of man with rabbit-virulent *Treponema paraluis-cuniculi*. *British Journal of Venereal Diseases*, 57(1), 7–10. <https://doi.org/10.1136/sti.57.1.7>
- Hallmaier-Wacker, L. K., Lueert, S., Roos, C., & Al., E. (2018). The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. *Sci Rep*, 8(6292). <https://doi.org/10.1038/s41598-018-24573-y>
- Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., & Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME Journal*, 5(10), 1571–1579. <https://doi.org/https://doi.org/10.1038/ismej.2011.41>
- Kallio, M. A., Tuimala, J. T., Hupponen, T., Klemelä, P., Gentile, M., Scheinin, I., ... Korpelainen, E. I. (2011). Chipster: User-friendly analysis software for microarray and other high-throughput data. *BMC Genomics*, 12. <https://doi.org/10.1186/1471-2164-12-507>
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, 41(1), e1–e1. <https://doi.org/https://doi.org/10.1093/nar/gks808>
- Klitgaard, K., Boye, M., Capion, N., & Jensen, T. K. (2008). Evidence of multiple *Treponema* phylotypes involved in bovine digital dermatitis as shown by 16S rRNA gene analysis and fluorescence in situ hybridization. *Journal of Clinical Microbiology*, 46(9), 3012–3020. <https://doi.org/10.1128/JCM.00670-08>
- Klitgaard, K., Bretó, A. F., Boye, M., & Jensen, T. K. (2013). Targeting the treponemal microbiome of digital dermatitis infections by high-resolution phylogenetic analyses and comparison with fluorescent in situ hybridization. *Journal of Clinical Microbiology*, 51(7), 2212–2219. <https://doi.org/10.1128/JCM.00320-13>
- Klitgaard, K., Nielsen, M. W., Ingerslev, H. C., Boye, M., & Jensen, T. K. (2014). Discovery of bovine digital dermatitis-associated *Treponema* spp. in the dairy herd environment by a targeted deep-sequencing

- approach. *Applied and Environmental Microbiology*, 80(14), 4427–4432.  
<https://doi.org/10.1128/AEM.00873-14>
- Knauf, S., Gogarten, J., Schuenemann, V., De Nys, H., D  x, A., Strouhal, M., ... Calvignac-Spencer, S. (2018). Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium *Treponema pallidum* subsp. *pertenue*. *Emerging Microbes & Infections*, 7(1).  
<https://doi.org/10.1038/s41426-018-0156-4>
- Kontturi, M., Junni, R., Simojoki, H., Malinen, E., Seuna, E., Klitgaard, K., ... Pelkonen, S. (2019). Bacterial species associated with interdigital phlegmon outbreaks in Finnish dairy herds. *BMC Veterinary Research*, 15(1), 1–12. <https://doi.org/10.1186/s12917-019-1788-x>
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Applied and Environmental Microbiology*, 79(17), 5112–5120. <https://doi.org/10.1128/AEM.01043-13>
- Krull, A. C., Shearer, J. K., Gorden, P. J., Cooper, V. L., Phillips, G. J., & Plummera, P. J. (2014). Deep sequencing analysis reveals temporal microbiota changes associated with development of bovine digital dermatitis. *Infection and Immunity*, 82(8), 3359–3373.  
<https://doi.org/10.1128/IAI.02077-14>
- Kuhnert, P., Brodard, I., Alsaad, M., Steiner, A., Stoffel, M. H., & Jores, J. (2020). *Treponema phagedenis* ( ex Noguchi 1912 ) Brumpt 1922 sp . nov ., nom . rev ., isolated from bovine digital dermatitis, 042818(ATCC 27087). <https://doi.org/10.1099/ijsem.0.004027>
- Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., ... Law, M. (2012). Comparison of next-generation sequencing systems. *Journal of Biomedicine and Biotechnology*, 2012. <https://doi.org/10.1155/2012/251364>
- Madden, T. (2002). Chapter 16 : The BLAST Sequence Analysis Tool. *The NCBI Handbook[Internet]*, (Md), 1–15.

- MADIGAN, M. T., BENDER, K. S., BUCKLEY, D. H., STAHL, D. A., & SATTLEY, W. M. (2019). *Brock biology of microorganisms*. New York: Pearson.
- Malmström, H., Storå, J., Dalén, L., Holmlund, G., & Götherström, A. (2005). Extensive human DNA contamination in extracts from ancient dog bones and teeth. *Molecular Biology and Evolution*, 22(10), 2040–2047. <https://doi.org/10.1093/molbev/msi195>
- McGovern, E., Waters, S. M., Blackshields, G., & McCabe, M. S. (2018). Evaluating established methods for Rumen 16S rRNA amplicon sequencing with mock microbial populations. *Frontiers in Microbiology*, 9(JUN). <https://doi.org/10.3389/fmicb.2018.01365>
- McMurdie, P. J., & Holmes, S. (2014). Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Computational Biology*, 10(4). <https://doi.org/10.1371/journal.pcbi.1003531>
- Neiman, M., Sundling, S., Grönberg, H., Hall, P., Czene, K., Lindberg, J., & Klevebring, D. (2012). Library Preparation and Multiplex Capture for Massive Parallel Sequencing Applications Made Efficient and Easy. *PLoS ONE*, 7(11). <https://doi.org/10.1371/journal.pone.0048616>
- Nielsen, M. W., Strube, M. L., Isbrand, A., Al-Medrasi, W. D. H. M., Boye, M., Jensen, T. K., & Klitgaard, K. (2016). Potential bacterial core species associated with digital dermatitis in cattle herds identified by molecular profiling of interdigital skin samples. *Veterinary Microbiology*, 186, 139–149. <https://doi.org/10.1016/j.vetmic.2016.03.003>
- Nordhoff, M., Taras, D., Macha, M., Tedin, K., Busse, H. J., & Wieler, L. H. (2005). *Treponema berlinense* sp. nov. and *Treponema porcinum* sp. nov., novel spirochaetes isolated from porcine faeces. *Int. Journal of Systematic and Evolutionary Microbiology*, 55, 1675–1680. <https://doi.org/10.1099/ijs.0.63388-0>
- Norris, S. J., Paster, B. J., & Smibert, R. M. (2015). *Bergey's manual of systematics of archaea and bacteria*. (W. B. Whitman, Ed.) (Vol. 410). Hoboken, NJ: Wiley. <https://doi.org/10.1002/9781118960608.gbm01249>



- Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, 18(5), 1403–1414.  
[https://doi.org/https://doi.org/10.1111/1462-2920.13023](https://doi.org/10.1111/1462-2920.13023)
- PEEK, S. F., & DIVERS, T. J. (2018). *Rebhun's Diseases of Dairy Cattle*. Saintt Louis: Elsevier. Retrieved from  
<https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5473511>
- Pokki, L., Helsingin yliopisto, M. t., University of Helsinki, F. o. A. a. F. & Helsingfors universitet, A. o. f. f. (2015). *Terveiden ja ajotulehdusta sairastavien nautojen sorkkavälin ihon bakteerilajien karakterisointi perinteisin mikrobiologisin menetelmin sekä MALDI-TOF-massaspektrometrillä*. Helsingin yliopisto.
- Polanski, A., & Kimmel, M. (2007). *Bioinformatics*. Berlin;New York: Springer c2007.
- Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. (2018). The Madness of Microbiome : Attempting To Find Consensus. *Applied and Environmental Microbiology*, 84(7), 1–12.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1). <https://doi.org/10.1093/nar/gks1219>
- Radolf, J.D., & Lukehart, S. . (Ed.). (2006). *Pathogenic Treponema: molecular and cell biology*. Horizon Scientific Press.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 2016(10), 1–22.  
<https://doi.org/10.7717/peerj.2584>
- Saarela, T., Helsingin yliopisto, E. t., University of Helsinki, F. o. V. M. & Helsingfors universitet, V. f. (2017). *Naudan sorkka-alueen ihotulehdus ja sen esiintyminen suomalaisissa pihattokarjoissa*. Helsingfors universitet.

- Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated - reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, 15(6), 1289–1303. <https://doi.org/10.1111/1755-0998.12402>
- Schrank, K. et al. (1999). *Treponema brennaborensis* sp. nov., a novel spirochaete isolated from a dairy cow suffering from digital dermatitis. *International Journal of Systematic Bacteriology*, 49, 43–50. Retrieved from <https://ijs.microbiologyresearch.org/content/journal/ijsem/10.1099/00207713-49-1-43?crawler=true&mimetype=application/pdf>
- Šmajs, D., Norris, S. J., & Weinstock, G. M. (2012). Genetic diversity in *Treponema pallidum*: Implications for pathogenesis, evolution and molecular diagnostics of syphilis and yaws. *Infection, Genetics and Evolution*, 12(2), 191–202. <https://doi.org/10.1016/j.meegid.2011.12.001>
- Sorkka-atlas. (2013). Pohjoismaiden. Retrieved from [http://www.faba.fi/wp-content/uploads/2018/12/Sorkkaterv\\_pm-sorkka-atlas\\_sf\\_claw\\_atlas\\_2013-09-02\\_webb.pdf](http://www.faba.fi/wp-content/uploads/2018/12/Sorkkaterv_pm-sorkka-atlas_sf_claw_atlas_2013-09-02_webb.pdf)
- Strub, S., Ploeg, J. R. Van Der, Nuss, K., Wyss, C., Luginbühl, A., & Steiner, A. (2007). Quantitation of *Gugenheimella bovis* and treponemes in bovine tissues related to digital dermatitis. *FEMS Microbiology Letters*, 269(1), 48–53. Retrieved from <https://doi.org/10.1111/j.1574-6968.2006.00604.x>
- Sullivan, L. E., Clegg, S. R., Angell, J. W., Newbrook, K., Blowey, R. W., Carter, S. D., ... Evans, N. J. (2015). High-level association of bovine digital dermatitis *Treponema* spp. with contagious ovine digital dermatitis lesions and presence of *Fusobacterium necrophorum* and *Dichelobacter nodosus*. *Journal of Clinical Microbiology*, 53(5), 1628–1638. <https://doi.org/10.1128/JCM.00180-15>
- Tizard, I. R. (2009). *Veterinary immunology: An introduction* (8th ed.). St Louis, Missouri.
- Zinicola, M., Lima, F., Lima, S., Machado, V., Gomez, M., Döpfer, D., Bicalho, R. (2015). Altered microbiomes in bovine digital

dermatitis lesions, and the gut as a pathogen reservoir. *PLoS ONE*, 10(3). <https://doi.org/10.1371/journal.pone.0120504>

**Liite 1. Chipster-ohjelmapaketin 16S rRNA-tuloslaskennan eri työvaiheet.**

<b>Työvaihe, engl.</b>	<b>Työvaihe</b>
	Tar-tiedostopaketin muodostaminen, yhdistää monta eri tiedostoa yhdeksi tiedostoksi.
	Sekvensoinnin laadun tarkkailu, FASTQC- tai MultiQC-raportin luonti
<a href="#">Trim primers and bar-codes and filter reads</a>	Alukkeiden ja indeksoinnin trimmaus sekä sekvenssien suodatus
<a href="#">Combine paired reads to contigs</a>	Yhdistetään parilliset sekvenssien lukutiedostot esitiedostoksi
<a href="#">Screen sequences for several criteria</a>	Erittäin pitkiä tai tuntemattomia emäksiä sisältävien sekvenssien poisto
Removal of identical sequences	Keskenään identtisten sekvenssien poistaminen
<a href="#">Align sequences to reference</a>	Sekvenssien linjaus referenssitietokantaan
<a href="#">Filter sequence alignment</a>	Linjauksen (alignment) suodattaminen eri parametreilla
<a href="#">Precluster aligned sequences</a>	Esiryhmittelee linjatut sekvenssit
<a href="#">Remove chimeric sequences</a>	Kimeeristen sekvenssien poisto
<a href="#">Classify sequences to units</a>	Luokittelee sekvenssit taksonomisiksi yksiköiksi eri kriteereiden mukaan
<a href="#">Produce count table and phenodata</a>	Luettelon muodostaminen bakteeritaksoneista
	Tilastolliset analyysit ja esim. Heatmap-esitys
	Rarefaktio, tulosten arpominen saaduista tuloksista

## Liite 2. Menetelmäkehityssekvensoinnin 19.11.2018 DNA-eristysparametreja.

Liite 2. Näytteet on ryhmitelty HBAK-numeroin, sorkan diagnoosi, tieto mahdollisesta kuulajahajotuksesta, Qubit BR tai HS mittaustulos, sekvenssilukumäärä

HBAK-numero	Sorkan diagnoosi	Kuulahajotus (beadbeating), [s]	DNA:n määrä, Qubit-tulos [ng/µl]	Sekvenssilukumäärä, Chipster-laskenta
509	M4	Ei kuulajahajotusta	6,47	183981
509B	M4	60	0,478	32707
510	M4	Ei kuulajahajotusta	24	107415
510B	M4	60	4,18	17502
511	M4.1	Ei kuulajahajotusta	2,62	156679
511B	M4.1	60	4,19	37927
513	M0	Ei kuulajahajotusta	7,33	972
513B	M0	180	2,78	4518
512	M4	Ei kuulajahajotusta	22,6	49582
512B	M4	60	2,96	138612
519	M4.1	Ei kuulajahajotusta	26,9	45545
519B	M4.1	60	4,32	57368
628	Känsiä M4.1	Ei kuulajahajotusta	2,86	8577
628B	Känsiä M4.1	180	2,91	5053
624	M4.1	Ei kuulajahajotusta	3,14	Alle 500
624B	M4.1	180	5,93	278752
625	M4.1	Ei kuulajahajotusta	26,3	82121
625B	M4.1	60	1,49	128550
517	M4.1	Ei kuulajahajotusta	21,6	138826
517B	M4.1	60	4,94	134446
RNALATER			Too low	0
PBS			Too low	0
PBS + MetaPol yzyme			Too low	0

**Liite 3. 29.5.2019 sekvensointiajon näytteet.**

Liite 3. Näytteet on ryhmittely HBAK-numeroin, rotu- ja poikimakerta, Kozichin työvuo, Chipster-laskennan perustella näytteestä saatu sekvenssilukumäärä.

<b>HBAK-numero</b>	<b>Rotu/ poikimakerta</b>	<b>Sorkan diagnoosi</b>	<b>Chipster laskennan sekvenssilukumäärä</b>
577	ay/1	Känsiä M0	6958
580	hol/1	Känsiä M0	6602
668	hol/2	M0	59958
664	ay/5	M0	90753
667	hol/2	M0	66751
627	hol/tieto puuttuu	M2	66841
626	hol/1	M2	106206
629	ay/tieto puuttuu	M2	56459
659	hol/2	M2	105912
652	hol/3	M2	114235

#### **Liite 4. Chipsterin 16S rRNA laskennan parametrit**

##### **Screen sequences for several criteria**

Maximum length = 275

Maximum number of ambiguous bases = 1

##### **Alignment**

silva.nr\_v132

##### **Screen sequences for several criteria**

Alignment start position = 13862

Alignment end position = 23444

Max. number of ambiguous bases = 1

Max. homopolymer length = 8

##### **Precluster aligned sequences**

Number of differences allowed = 1

##### **Remove chimeric sequences**

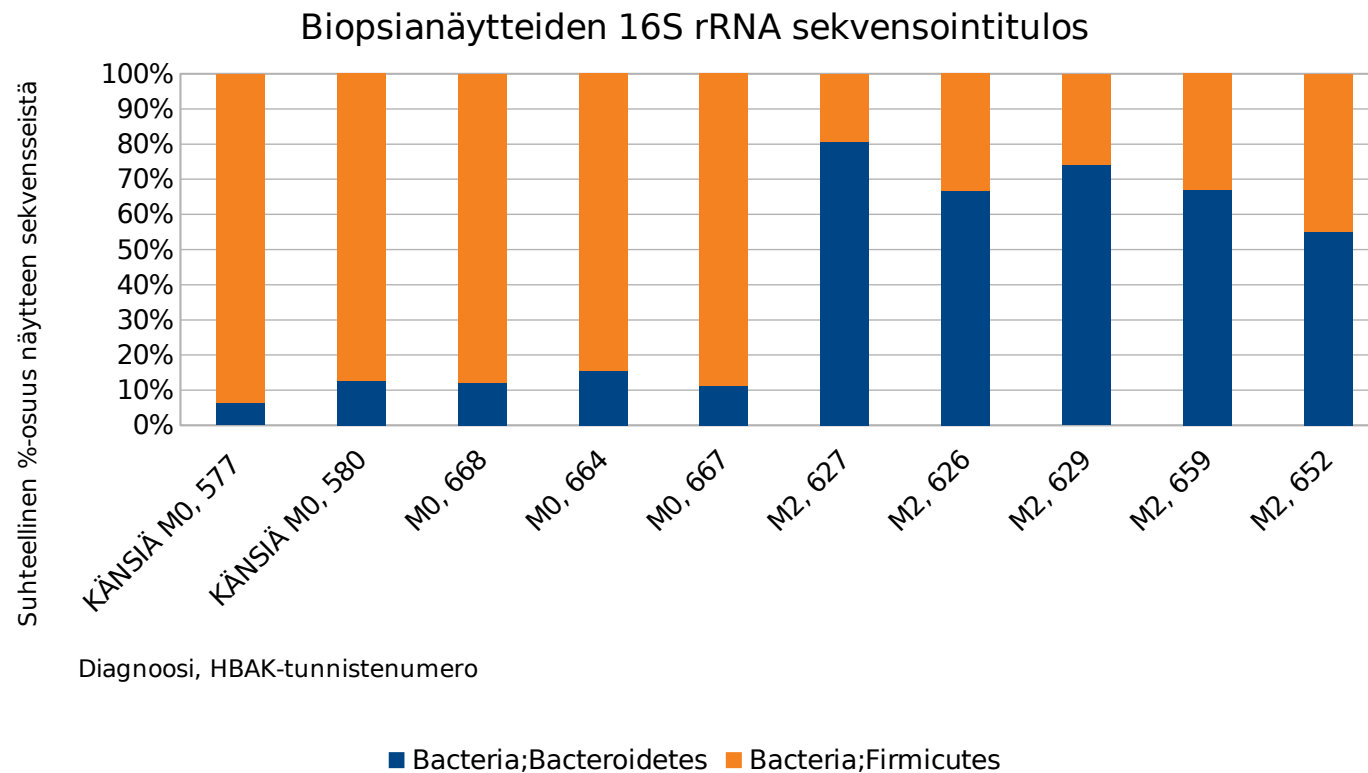
dereplicate=true

16S rRNA Silva gold bacteria

##### **Classify sequences to taxonomic units**

Bootstrap = 1000

## Liite 5. *Bacteroidetes/Firmicutes*-suhteen tarkastelu.



Liite 5. 16S rRNA sekvensointitulos, *Bacteroidetes/Firmicutes*-suhteen tarkastelu. Chipster-laskenta, DNA-eristysmenetelmä QuickPick SML gDNA -kitillä., alukkeet ja indeksointi Kozich ym. 2013 mukaisesti, MiSeq, V2 2x250 bp, ajopäivämäärä oli 29.5.2019.